THE JOURNAL IS INCLUDED IN THE LIST OF THE LEADING JOURNALS AND EDITIONS OF THE HIGHER ATTESTATION COMMISSION (VAK)

ISSN 1729-2646
e-ISSN 2500-3909

THE JOURNAL PROMOTER:
"Journal "Reliability" Ltd

It is registered in the Russian Ministry of Press, Broadcasting and Mass Communications. Registration certificate ПИ 77-9782, September, 11, 2001.

Official organ of the Russian Academy of Reliability

Publisher of the journal
LLC Journal "Dependability"
Director
Dubrovskaya A.Z.
The address: 109029, Moscow,
Str. Nizhegorodskaya, 27,
Building 1, office 209
Ltd Journal "Dependability"
www.dependability.ru
Printed by JSC "Regional printing house,
Printing place" 432049, Ulyanovsk,
Pushkarev str., 27. Circulation: 500 copies.
Printing order
Papers are reviewed. Signed print ,
Volume , Format 60x90/8, Paper gloss

Papers are reviewed.
Papers are published in author's edition. The opinion of members of the editorial board may not coincide with the point of view of authors' publications. The reprint of materials is granted only with the written permission of the editorial board. Manuscripts are not returned.

THE JOURNAL IS PUBLISHED WITH THE PARTICIPATION AND SUPPORT OF THE JOINT-STOCK COMPANY «RESEARCH AND DESIGN INSTITUTE OF INFORMATISATION, AUTOMATION AND COMMUNICATION ON RAILWAY TRANSPORT» (JSC «NIIAS») AND LLC PUBLISHING HOUSE «TECHNOLOGY»

CONTENTS

Structural dependability. Theory and practice

Shubinsky I.B., Zamyshliaev A.M., Papić L. Adaptive dependability of information management systems 3

Fedukhin A.V., Cespedes Garcia N.V. On the matter of evaluation of the variation coefficient of the time to failure based on low level quantiles 10

Kitaev S.V., Baikov I.R., Smorodova O.V. Set of indicators for dependability evaluation of gas compression units 16

Bazhenov Yu.V., Bazhenov M.Yu. Research of operational dependability of automotive engines 22

Pokhabov Yu.P. What should mean dependability calculation of unique highly vital systems with regards to single-use mechanisms of spacecraft 28

Functional safety. Theory and practice

Klimov S.M., Sosnovsky Yu.V. Method of assessing the protection of computer-based control systems under information technology interference 36

Zviagin V.I., Ptushkin A.I., Trudov A.V. Risk as one of the properties of decisions taken under uncertainty 45

Account

Kuzmina N.A. Ensuring an efficient transportation infrastructure security system by means of solutions that enable detection of intrusions into protected areas 51

Gnedenko Forum 56

Original article
Dependability, vol. 18 no.4, 2018 DOI: 10.21683/1729-2646-2018-18-4-3-9

Adaptive dependability of information management systems

Igor B. Shubinsky, *JSC NIIAS, Moscow, Russia*
Aleksey M. Zamyshliaev, *JSC NIIAS, Moscow, Russia*
Ljubi a P. Papi, *Research Center of Dependability and Quality Management, Prijevor, Serbia*

Igor B. Shubinsky

Aleksey M. Zamyshliaev

Ljubi a P. Papi

Abstract. *The paper examines the reliability of an information management system as its ability to provide the required services that can be justifiably trusted. It is assumed that the system functions without an operator. The aim is to ensure the dependability of a multimodule control system, when the problem-solving results are affected by failures, faults and errors of problem-solution by the system's computation modules (CMs). Conventional fault tolerance methods do not provide the desired effect, as even under infinite structural redundancy yet real capabilities of on-line detection of CM failures or faults the system's dependability is significantly lower than expected. The paper proposes and evaluates the methods of adaptive dependability. They are to ensure the observability of control systems under limited capabilities of component CM operability supervision, as well as achieving the required levels of dependability of information management systems in cases of insignificant float time and structural redundancy. These goals are achieved through active (and automatic) reassignment of the available computational resources for on-line information processing. The methods of adaptive dependability enable – with no interruption of computational processes and while solving real-world problems – timely automatic detection and elimination of failures, faults of CMs and errors in the solution of specified problems through on-line localization of faulty modules and subsequent automatic reconfiguration of the system with the elimination of such modules from operation.*

Keywords: *computation modules, dependability, adaptive protection, failures, faults, errors in performance of designated tasks, automatic system reconfiguration, control, allowed time of interruption of operation, time redundancy, protection cycles and beats.*

For citation: *Shubinsky IB, Zamyshliaev AM, Papi L. Adaptive dependability of information management systems. Dependability 2018;4: 3-9. DOI: 10.21683/1729-2646-2018-18-4-3-9*

1. Introduction

1.1. Dependability of information management systems

The matters related to ensuring the reliability of information technology are the main focus of all experts directly or indirectly involved in its development, manufacture and operation. Over the years of digital technology development the failure rate of the basic components decreased by six orders of magnitude. An information system includes thousands of digital elements each of which is a hardware and software device performing a multitude of functions.

Now, the key problem of ensuring the reliability of an information system is the faultless performance of the assigned functional tasks that, in technical terms, are implemented by information processes. The relevance of this problem is due to the fact that error rate in the operation of an information system and the associated functional failure rate significantly exceed the failure rate of digital technology, while the functional failures themselves may be critical to the environment and controlled objects [1, 2, etc.]

Due to that some researchers assume that reliability of information technology performance should be studied as the ability of an information system to deliver service that can be trusted. The *service* delivered by a system is its properties or behavior as it is perceived by its *user*. In the interpretation of this paper's authors a service that can be trusted is perceived as *overall reliability* [3].

In the mentioned paper [3] the following concepts are used:

correct service is delivered when the service implements the system function(s);

system *failure* is an event that occurs when the delivered service deviates from correct service, i.e. failure is a transition from correct service to *incorrect service*, when the system function is not implemented.

The development of this approach is reflected in the research paper of the Working Group 10.4 of the International Federation for Information Processing [4]. However, instead of the term "overall reliability" the group's experts introduce the term *"dependability"* that in this paper is considered as the "trustworthiness of a computing system which allows reliance to be *justifiably* placed on the service it delivers." Service is a form of activities that do not create a new material product, but change the quality of an existing previously created product. The delivery of service itself creates the desired result [5]. Explicitly, dependability is a property of the *service* and depends on the system's utilization.

1.2. Limitations of the conventional methods of ensuring dependability of control systems

The delivery of service to user with the given level of guaranteed quality is performed with the help of a technical system and is an action, process required for the implementation of the service delivery function. Here we imply the combination of hardware, software and human operator of the information system. Hereinafter, we assume that the control system automatically performs the specified functions without the involvement of the human operator. Consequently, ensuring a high level of system dependability requires prior achievement of even higher level of hardware (product) and software components reliability. The products are an object or a set of objects manufactured at an enterprise. The classic (structural) reliability theory examined the processes of *products* (system, element) failures and recoveries. In [1, 6], it is shown that even under arbitrarily large redundancy it does not appear to be possible to achieve a high level of product reliability. The object of the research was the reliability model of a redundant object with partial redundancy composed of one primary and an infinite number of same-type backup devices. The following is assumed:

• The components' lifetime duration is a random value and is described with a service life distribution that meets the following conditions:

- the times of outage of each of the backup components are statistically independent from each other;

- all the backup components have an identical exponential distribution of service life.

• The system of these random values is an ordinary recovery process.

• The supervision and commutation facilities to the backup devices are perfectly dependable.

• The switch time is negligibly small.

Under the given premises, the limit probability of no failure of a redundant group is defined as $P_L(t) \le \sum_{n=0}^{\infty} P(n,t) \cdot y^n$, where $P(n,t)$ is the distribution of the resultant number of the time intervals between the replacements of failed devices of a specific facility, that before the failure performed the functions of the main element; is $\gamma = P\{v \le \tau_A\} = \int_0^{\tau_A} f_v(t)dt = F_v(\tau_A)$ the probability of correct and timely detection of failure and backup switching, v is the random device failure duration, τ_A is the allowed duration of system outage (for control systems this time is comparable with the duration of control cycle); $f_v(t)$ is the density function of failure duration in the system.

Under the above assumptions [7] established that the mean time to failure of a redundant object with partial redundancy composed of one primary and an infinite number of same-type backup devices does not exceed the level defined by formula (1) on the assumption of simple device failure or fault flow

$$T_{FS} \le 1/\lambda(1-\gamma) \quad (1)$$

where λ is the failure rate of one device.

In [7], it is established that the expected increase of the mean time to failure of the initial device due to multiple redundancy with recovery can not be more than 2…10 times even under a very high probability of correct and timely detection of failure and backup switching $0.8 < \gamma \le 0.9$.

Taking into consideration that the system's software is also executed with errors and more often with faults [8, 9, 10, etc.], the achievement of a high level of system dependability by means of conventional methods should not be expected even under condition of heavy investment into system redundancy.

2. Definition of the problem of adaptive dependability

It is required to ensure the specified high level of information management system dependability ***without introducing large*** structural, time, functional and other redundancy by means of:

- dependability management based on the results of evaluation of the ***correctness*** of system tasks performance, not the rate of failures and recovery;
- use of ***natural*** time redundancy that persists in many systems within the control cycle;
- ***adaptation*** of the system to erroneous results of system tasks performance with dynamic rearrangement of the system and parallel performance of tasks with beat-to-beat comparison of results;
- ***priority*** handling of the most important tasks in order to ensure their higher dependability.

The ideas and principles of adaptive dependability have much in common with the concept of active protection (AP) that we set forth in [11]. They can be briefly described as follows:

- the duration of all cycles of the information processing divides into certain constant or random time intervals that shall be further called beats, within each of which the specified set of software modules are executed and hold points are formed;
- the whole set of the constituent computation modules (CMs) of an information system is divided into two compound sets: the computing environment, i.e. a set of m same-type CMs; the protective environment, i.e. a set of $k \le m$ same-type CMs redundant in terms of the specified tasks;
- dynamic rearrangement of control system modules is carried out at every second beat for the organization of parallel information processing;
- beat-by-beat virtual redundancy by means of parallel solution of all specified m tasks at the primary CMs provided there is at least one operational redundant CM;
- minimal system configuration must include not less than $m = 2$ main and one redundant CM for detection of erroneous result in the solved task, classification and localization of malfunctions;
- synthesis of adaptive dependability (AD) is based on the selection of the value of beat duration τ, under which within the allowed duration of outage the error in the task solution must – with the specified level of assurance – be detected and eliminated through the localization of the error source CM and its swapping for an operable redundant CM.

3. Organization of systems with adaptive dependability

Different disciplines can be suggested for the practical implementation of ideas and methods of AD organization. In this paper, two disciplines are examined, i.e. ***D1*** and ***D2***.

D1. A system with one-beat restart containing m main and one controlling CM, non-priority control, no reassignment of modules. In the case of failure of one of a pair of CMs repeated calculation with the previous operands is performed. Matching results in the next beat eliminates the possibility of failure of modules, the failure has been eliminated, the hold point of assignment of the first CM in the i-th protection cycle is updated. If CM fault is detected by own control facilities, the hold point is naturally updated based on the data of the first main CM. A failure of one of the pair of CMs is detected by means of a restart for one AP beat. If, in the process of solution of the same part of a task, over two beats the results of the operation of a pair of same-type CMs do not match twice, the hold point is not updated until joint operation of the controlling CM with the next main (the third in this example) CM. In case of matching results for this last pair the decision is made regarding the failure of the previous main CM (the second one in this example), the hold point for the second CM is updated based on the data of the controlling module that now performs the role of the second main CM. If in three adjacent protection beats the results do not match, the decision is made regarding the failure of the controlling CM and the system may for some time operate without protection, if there is no operable backup module.

Thus, relative to discipline ***D1*** the parameters A, b and x_E are characterized by the following: number of beats in the protection cycle $A = m$; number of main CM failure or fault decision-making beats $b = 2$; number of beats for recovery of computation process from the last hold point $x_E \le m + 2$.

Table 1

Number of beat	Numbers of primary CMs	Number of controlling CM	Pairs of controlled CMs	Reassigned CMs	CM polling rate		
					2	5	1, 3, 4, 6, 7, 8
1	1 8 3 4 5 6 7	2	**2**–7	8–2	**4**	**2**	**1**
2	1 8 3 4 5 6 7	2	**2**–3	8–2			
3	1 2 3 4 5 6 7	8	8–**5**	-			
4	8 2 3 4 5 6 7	1	1–**2**	8–1			
5	1 2 3 8 5 6 7	4	4–**2**	8–4			
6	1 2 3 4 8 6 7	5	**5**–6	8–5			

D2. A system with restart and CM reassignment containing m main and one controlling module. The organization of detection and elimination of malfunctions is the same as in discipline ***D1.*** *CM reassignment* is required in order to shorten the protection cycle, when the number of main CMs is significantly higher than that of the backup modules. The point of reassignment consists in the fact that in specific beats CMs are redistributed between the computing and protective environments. For the time of a beat some modules of the protective environment are assigned the functions of main modules and vice versa. This eliminates the inherent weakness of methods of controlling CM fixation, when the modules of the computational environment are controlled much less frequently that those of the protective environment. Indeed, in all cases of fixation the modules of the protective environment within the AP cycle take part in all pairs of controlled CMs, whereas the modules of the computational environment take part in just one pair, or somewhat more frequently, if in each protection beat two and more CM pairs are formed.

Thus, relative to discipline ***D2*** parameters A b and x_E are as follows: $A = \text{int}\left(\frac{m+1}{2}\right)$, $b = 2$, $x_E = \text{int}(\frac{m+5}{2})$.

Organization of ***priority control*** of the control system's ability to correctly solve the specified tasks allows significantly increasing the level of its dependability in terms of priority tasks. Priority control is organized by means of CM reassignment. However, the intention is different. Whereas the reassignment of modules aimed to equalize the frequency of controls of main and backup CMs, priority control aims to increase the frequency of control of the modules most significant in terms of the specified tasks.

Let us illustrate the feasibility of systems with two modules identified as priority (Table 1). It is assumed that the first identified module (zero priority) is controlled in the AP cycle with the assigned maximum frequency, the second one (first priority) is controlled with increased frequency, yet it is lower than with the zero priority module. The remaining CMs in the system are controlled with an equal frequency that is yet lower than that of the priority modules. Let $m = 7$, $k = 1$ ($m + k = 8$), zero priority is given to module 2, while first priority is given to module 5. Let us stipulate that in the AD cycle module 2 was controlled in four beats, module 5 was controlled in two beats, while the remaining modules 1, 3, 4, 6, 7 and 8 were controlled in one beat. The solution of this problem is shown in Table 1.

The following results were obtained. AD cycle $A = 6$ beats, CM are reassigned four times, module 2 is controlled in two beats out of three adjacent ones, while module 5 is controlled at every third beat. The duration of AD cycle increased 1.5 times compared to uniform CM reassignment, since in that case the duration of cycle would be $A = (m + k)/2k = 4$. This is natural, since the reduction of time intervals between the controls of some CMs is possible at the expense of longer intervals between the controls of non-priority CMs. Solving such AP problems should involve reasonable trade-offs. This applies fully to the selection of the method of CM fixation or reassignment. In the first case AD management is simpler, in the second case control cycle is shorter. Reassignment of CM is more preferable in case of very low values of allowed duration of outage, although AB management is somewhat more complicated. Under less strict time restrictions AG should be attempted to be implemented by means of fixation of controlled CMs.

4. The efficiency of the methods of adaptive dependability of control systems

The efficiency of adaptive dependability is evaluated based on the indicator of probability of successful adaptation of an information management system to failures, faults, software errors. The adaptation will be successful if as the result of the actions performed as part of the protection algorithms the duration of the specified malfunctions is less or equal to the allowed value, which enables the elimination of erroneous results in the control process. The allowed value means the control cycle, i.e. the time within which the detection and elimination of system malfunction will not cause subsequent erroneous control. Since the elimination time for each protection discipline is a constant number of AD beats, it will suffice to compare the duration of malfunction detection with the allowed detection time.

Let us perform the verification of the efficiency of adaptive dependability for the following types of protection organization conditions.

The tasks of information processing are divided into equal parts (beats) τ, with the duration of a beat being much shorter than the duration of the task. The tasks are solved with random time intervals υ_2, however the duration of task solutions υ_1 are much longer than the duration of pauses, i.e. $\upsilon_1 >> \upsilon_2$. That allows dividing the task into protection beats (e.g. for generality, random duration beats). Additionally, it is taken into consideration that the allowed outage of the system is a constant value τ_A. It is assumed that there are no simultaneous failures or faults of the operable and controlled CM that is verified within the current beat. The duration of the beat is defined by the duration of execution within the beat of a group of functionally complete software modules. Since all CMs that execute software modules are same-type, the order of distribution of the software modules per CM operation beats is common for all CMs. This allows adopting the distributions $F_\upsilon(t)$ of beat duration as identical for all CMs.

It is required to establish the probability of the system's successful adaptation to failures:

$$\beta = P\{\nu \le \tau_A - t_E\} = \int_0^{\tau_A - t_E} f_\nu(t)dt \qquad (2)$$

where $f_\nu(t)$ is the density function of the time ν of a dormant fault's existence in the system.

In order to find the functions of density $f_\nu(t)$ and probability of successful adaptation to failures β in general, the following parameters are used:

• distribution functions and characteristics of protection time intervals, i.e. beat duration, allowed time of system outage and time of elimination of detected failure (τ_A and t_E respectively);

• parameters of the adopted AP discipline: $A, b, t, x_E = t_E/\tau$.

The time of connection of the controlling CM to the next main CM consists of the random duration of beat υ and wait time ψ from the moment of completion of the parallel operation with the previous CM to the moment of the beginning of the next operation beat of the next CM. $\psi \le \upsilon$ and during the wait time the memory of the controlling CM is loaded with commands and operands of the next main module.

For each time density function ν let us preliminarily set the total time density function ψ + υ. In the Laplace domain it is as follows

$$f_\varsigma(s) = \phi_\psi(s) * f_\upsilon(s),$$

where $\phi_\psi(s)$ is the portrayal of the distribution density of wait time ψ, while $f_\upsilon(s)$ is the portrayal of the distribution density of the duration of the AP beat.

Let us assume that between the occurrence of a dormant failure of CM and the moment the controlling CM connects to it x beats elapsed. Then, the conditional probability of $x < X$, where X = 0, 1, …, A, …, can be found using the appropriate Laplace transformation

$$f_x(s) = [f_\xi(s)]^x.$$

Due to the equally likely possibility of failure of any CM that are not protected during the current beat, it can be assumed that the integer random variable x is uniformly distributed over the number range 1, 2, …, A–1. Out of this, the distribution density of the number of beats of malfunction existence within the system is identified using the following formula:

$$f(x) = \sum_{i=1}^{A-1} \frac{\delta(x-i)}{A-1}, \qquad (3)$$

where $\delta(x)$ is the delta function of parameter x.

The total duration of failure existence until its detection is the sum of time x(υ+ψ) and time b(υ+ψ) from the moment of detection of the fact of malfunction to the localization of the failed CM in accordance with the chosen AP discipline.

The density function of random value x(υ+ψ)=θ in the Laplace image is depicted as follows according to the total probability formula.

$$f_\theta(s) = \sum_{i=1}^{A-1} \frac{1}{A-1} \left(f_\xi(s)\right)^i.$$

The density function of random value (x+b)· (υ+ψ) in the Laplace image is calculated as

$$f_\nu(s) = f_\theta(s) * (f_\upsilon(s))^b = \frac{1}{A-1} \sum_{i=1}^{A-1} \left(f_\xi(s)\right)^{i+b} \qquad (4)$$

The next step in the identification of the probability of successful adaptation to failures of a system with AP design under consideration consists in developing function $f_\xi(s)$ in the above formula, that in the Laplace image is the density function of the sum of beat duration and time delay of controlling CM connection to the main module within the beat ($\xi = \upsilon + \psi \le 2\upsilon$).

Using experimental data [2] let us take the distributions of random beat durations υ as an Erlang distribution of the a-th order with the density function $f_\upsilon(t) = \frac{\rho(\rho \cdot t)^a}{a!} e^{-\rho \cdot t}$ that in the Laplace image are as follows:

$$f_\upsilon(s) = \left(\frac{\rho}{\rho + s}\right)^{a+1},$$

where ρ is the Erlang distribution parameter (number of events per unit of time).

According to [12], the density function of wait time ψ (in our case, the time of controlling CM connection to the main CM) in the Laplace image is as follows:

$$\phi_\psi(s) = \frac{\rho}{(a+1)s}\left[1 - \left(\frac{\rho}{\rho + s}\right)^{a+1}\right].$$

Consequently, in formula (4) density function $f_\xi(s)$ equals to

$$f_\xi(s) = f_\upsilon(s) \cdot \phi_\psi(s) = \left(\frac{\rho}{\rho + s}\right)^{a+1} \cdot \frac{\rho}{(a+1)s}\left[1 - \left(\frac{\rho}{\rho + s}\right)^{a+1}\right].$$

By substituting this formula into formula (3.4) we deduce that

$$f_\nu(s) = \frac{1}{A-1} \sum_{i=1}^{A-1} \left\{\left(\frac{\rho}{\rho + s}\right)^{a+1} \frac{\rho}{(a+1)s}\left[1 - \left(\frac{\rho}{\rho + s}\right)^{a+1}\right]\right\}^{i+b}.$$

By moving from the image to the original under a constant value of the allowed outage time and using formula (3) we identify the probability of successful adaptation to failures of a system with AD

$$\beta = 1 - \frac{e^{-(a+1)\cdot x_A^*}}{A-1} \sum_{i=1}^{A-1} \left(\frac{1}{a+1}\right)^{i+b} \sum_{|\bar{n}|=i+b} \frac{(i+b)!}{\bar{n}!} \cdot \sum_{k=0}^{\eta} \frac{((a+1)x_A^*)^k}{k!} \qquad (5)$$

where $\bar{n}! = n_0! \cdot n_1! \cdot \ldots \cdot n_a!$; $|\bar{n}| = n_0 + n_1 + \ldots + n_a$; $x_A^* = \tau_A^*/\tau$; $t_A^* = \tau_A - t_E$;

$$\eta = (a+2)(i+b) + \sum_{j=1}^{a} j n_j + 1.$$

In the special case a = 0 (exponential distribution of beat duration) the following formula for the probability of the system's successful adaptation to failures is true $\beta = 1 - \frac{e^{-x_A^*}}{A-1} \sum_{i=1}^{A-1} \sum_{k=0}^{2(i+b)+1} \frac{(x_A^*)^k}{k!}$, as in this case $\bar{n}! = 1$, while $|\bar{n}| = 0$.

Using expression (5) let us analyze the dependence of the probability of successful adaptation of a system with AP from the allowed number of outage beats, number m of main modules and subject to the above examined disciplines ***D1*** and ***D2***.

Figure 1 shows the dependences $\beta = f(x_A)$ under $a \ge 2$ in respect to disciplines ***D2*** (solid lines) and ***D1*** (dotted lines).

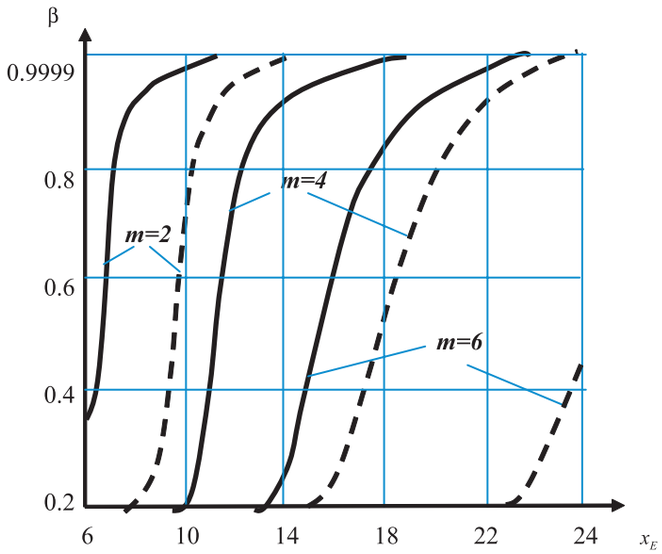


Figure 1. Dependences of the probability of successful adaptation of a system with random protection beats to failures depending on the allowed number x_E of protection beats and number m of main CMs

Beginning from the second order of the Erlang distribution of beat durations and higher the results of such dependences are practically identical. This shows that disciplines similar to ***D2*** have the highest speed of adaptation to CM failure. These disciplines react to the errors in the task solution results about a few beats quicker than the discipline of class ***D1***. The advantages of the above disciplines increase with the number of main computation modules.

At the same time, ABs with random beat duration are much more inertial than ABs with constant beat duration. Thus, even under the minimal for AD number of main modules m = 2 the time of detection and elimination of CM failure increases 1.5 – 2 times. Since for many control system architectures and associated computational processes it does not appear to be possible to provide AD with constant beats, additional opportunities of increasing in the speed of adaptation of system with AD to failures of component CMs should be found. For instance, such opportunity exists if built-in control of main CMs is also used that can accelerate the detection and elimination of CM failures in systems with AD.

5. Conclusion

Limited capabilities to ensure redundancy, on-line detection of failures, faults, errors of information process performance, as well as the limited capabilities of the hardware and software system require the development of unconventional technological solutions to ensure dependability of information management systems. One of them is the adaptive dependability technology proposed in this paper. Essentially, it consists in the active use of natural time and structural redundancy and active (and automatic) reassignment of available processing resources not only for real-time information processing, but also for observability of the system under limited supervision facilities. Adaptive dependability is intended for enabling the required levels of dependability of information management systems under insignificant time margin, limited efficiency of component processing modules fault detection facilities, as well as under the condition of the amount of redundant equipment not exceeding the amount of primary equipment. Adaptive protection provides viable opportunities of achieving a much higher level of dependability compared to conventional redundancy methods. The adaptive dependability technology enables – under restricted time while solving real-world problems – timely automatic detection and elimination of failures and faults through on-line localization of faulty modules and subsequent automatic reconfiguration of the system with the elimination of such modules from operation. At the same time, this technology is geared towards multimodule systems and is not adapted for systems of information storage and display, documentation, power supply of information management systems.

References

[1]. Shubinsky IB. Nadiozhnye otkazoustoychivye informatsionnye sistemy. Metody sinteza [Dependable failsafe information systems. Synthesis methods]. Moscow: Dependability Journal; 2016 [in Russian].

[2]. Kirpichnikov AP, Vasiliev SN. Particular characteristics of today's microelectronics and matters of highly dependable and secure control systems design. Dependability 2017;3:10-16.

[3]. Avizienis A, Laprie J-C and Randell B. Dependability of computer systems. Fundamental concepts, terminology and examples. Technical report. LAAS – CNRS; October, 2000.

[4]. Rus I, Komi-Sirvio S, Costa P. Computer program with insurance of high reliability. Technical report. IFIP WG-10.4; March, 2008.

[5]. Borisov AB. Bolshoy ekonomicheskiy slovar [Large economic dictionary]. Moscow: Knizhny mir; 2003 [in Russian].

[6]. Shäbe H, Shubinsky IB. Limit reliability of structural redundancy. Dependability 2016;1:9-13.

[7]. Shubinsky IB. Methods of software functional dependability assurance. Dependability 2014;4:95-101.

[8]. Potapov IV. Issues of software systems dependability. Dependability 2015;1:58-61.

[9]. Shubinsky IB, Schäbe H. A systematic approach to protection against glitches. Dependability 2014;3:103-107.

[10]. Shubinsky IB, Shäbe H. On the definition of functional reliability. In: Steenbergen et al., editors. Proceedings of the ESREL 2013, Safety, Reliability and Risk Analysis: Beyond the Horizon. London (UK): Taylor & Francis Group; 2014. pp. 3021-3027. ISBN 978-1-138-00123-7.

[11]. Shubinsky IB. Adaptive fault tolerance in real-time information systems. Life Cycle Engineering and Management. In: Proceedings of ICDQM-2016. Prijevor (Serbia); 29-30 June 2016. pp.3-14.

[12]. Gnedenko BV, Kovalenko IN. Vvedenie v teoriyu massovogo obsluzhivaniya [Introduction into the waiting theory]. Kiev: Nauka; 1963 [in Russian].

About the authors

Igor B. Shubinsky, Doctor of Engineering, Professor, Deputy Head of Integrated Research and Development Unit, JSC NIIAS, Moscow, Russia, phone: +7 (495) 786 68 57, e-mail: igor-shubinsky@yandex.ru

Aleksey M. Zamyshiaev, Doctor of Engineering, Deputy Director General, JSC NIIAS, Moscow, Russia, phone: +7 (495) 967 77 02, e-mail: A.Zamyshlaev@vniias.ru

Ljubiša Papić, DR.SC in Engineering, Professor, Director, Research Center of Dependability and Quality Management, Prijevor, Serbia

Received on 26.08.2018

Dependability, vol. 18 no.4, 2018
Original article
DOI: 10.21683/1729-2646-2018-18-4-10-15

On the matter of evaluation of the variation coefficient of the time to failure based on low-level quantiles

Alexander V. Fedukhin, *Institute of Mathematical Machines and Systems Problems, NAS of Ukraine*
Natalia V. Cespedes Garcia, *Institute of Mathematical Machines and Systems Problems, NAS of Ukraine*

Alexander V. Fedukhin

Natalia V. Cespedes Garcia

Abstract. *In the context of various tasks related to dependability estimation of systems by probabilistic physical methods the most important a priori information that ensures effective solutions is the information on the variation coefficient of the time to failure. Given the low failure statistics, the estimation of the variation coefficient of the time to failure is complicated due to significant sample censoring. In these cases, methods of variation coefficient evaluation with additional a priori information and the method of quantiles are used. The solution of a number of dependability-related tasks that require taking into consideration various failure distributions is significantly simplified if the functions of such distributions are tabulated in the relative operation time and variation coefficient parameters. An effective solution of dependability-related tasks with the use of tables of DN distribution function was first proposed for the parametrization of distribution in parameters x and v, where x is the scale parameter, relative operation time x = at; v is the shape parameter, variation coefficient v = V; a is the average degradation rate. That allowed performing tabulation out of real time, simplifying function tabulation and its use in a number of dependability-related tasks by method of quantiles. The paper analyzed the effectiveness of the method of quantiles in the estimation of the variation coefficient of the time to failure, that is at the same time the shape parameter of the DN distribution, under scarce failure statistics and based on it proposes a new, more effective, method. The method of estimation of the variation coefficient using low and ultralow-level quantiles is based on the behaviour analysis of function $a_i = f(t)$ obtained using the method of quantiles. It is considered that the best choice of the a priori value of v is a choice under which the dependence graph $a_i = f(t)$ is most accurately described by a straight horizontal line, which is in complete compliance with the hypothesis of constant degradation rate accepted in the context of DN distribution formalization. In cases when the dependence graph $a_i = f(t)$ does not easily allow concluding on the best choice of the a priori value v (it is especially difficult to make a choice based on the statistics of first failures), the following formal criterion can be used: the most acceptable a priori value of the shape parameter v lies within the range of values, where the sign of the trend of the average degradation rate (h) in graph $a_i = f(t)$ changes. Studies have established that the most significant errors in the estimation of the variation coefficient are associated with first failures. When processing the results of dependability tests it is assumed the first failures in a sample have the lowest information weight, as their occurrence is due to serious defects not detected by final quality inspection of products. The first failures normally "fall out" of the overall statistical pattern, and it is recommended to omit them from further analysis. The proposed method of estimation of the variation coefficient of the time to failure based on ultralow-level quantiles enables – in the context of limited failure statistics, when other methods are inefficient – for sufficiently accurate identification of not only the variation coefficient of the time to failure and DN distribution parameters, but also make conclusions regarding the feasibility and legitimacy of equalization (description) of the considered sample using this diffusion distribution, i.e. it can be used as a kind of criterion of compliance of the empirical failure distribution under consideration with the chosen theoretical dependability model. The described process of finding the truest values of the variation coefficient of the time to failure using the formal criterion can be computerized.*

Keywords: *method of quantiles, variation coefficient, low and ultralow-level quantiles, DN distribution.*

For citation: *Fedukhin AV, Cespedes Garcia NV. On the matter of evaluation of the variation coefficient of the time to failure based on low-level quantiles. Dependability 2018;18(4); 10-15. DOI: 10.21683/1729-2646-2018-18-4-10-15*

1. Introduction

In the context of various tasks related to dependability estimation of systems by probabilistic physical methods [1] the most important a priori information that ensures effective solutions is the information on the variation coefficient of the time to failure. Given the low failure statistics the estimation of the variation coefficient of the time to failure is complicated due to significant sample censoring. In these cases methods of variation coefficient evaluation with additional a priori information [2-4] and method of quantiles are used [1].

The paper analyzed the effectiveness of the method of quantiles in the estimation of the variation coefficient of the time to failure (shape parameter of the *DN* distribution [5-8]) under scarce failure statistics and based on it proposes a new, more effective, method.

2. Method of quantiles

If the a priori value of the shape parameter v is known, whose consistent estimate is the variation coefficient of the degradation process V, the scale parameter of the *DN* distribution, i.e. the average degradation rate a, can be identified by solving equation [1]:

$$\Phi\left(\frac{at_\gamma - 1}{v\sqrt{at_\gamma}}\right) + \exp(2v^{-2})\Phi\left(-\frac{at_\gamma + 1}{v\sqrt{at_\gamma}}\right) = \hat{\gamma}, \qquad (1)$$

where $\hat{\gamma} = r / N$ is the quantile calculated based on the ratio of the number of failures r to the sample size N submitted to tests; t_γ is the time of occurrence of the r-th failure.

The solution of a number of dependability-related tasks that require taking into consideration various failure distributions is significantly simplified if the functions of such distributions are tabulated. An effective solution of dependability-related tasks with the use of tables of the *DN* distribution function was first proposed in [9], where the *DN* distribution function was parametrized and tabulated in the x and v parameters. The use of the relative operation time $at = x$ as the distribution parameter allowed performing tabulation out of real time, simplifying function tabulation and its use in a number of dependability-related tasks by method of quantiles.

$$\Phi\left(\frac{x_\gamma - 1}{v\sqrt{x_\gamma}}\right) + \exp(2v^{-2})\Phi\left(-\frac{x_\gamma + 1}{v\sqrt{x_\gamma}}\right) = \hat{\gamma}, \qquad (2)$$

where $x_\gamma = at_\gamma$.

With the use of *DN* distribution tables [1] and input data on $\hat{\gamma}$ and v the value of x_γ is identified, then formula $a = \frac{x_\gamma}{t_\gamma}$ is used to calculate the value of average degradation rate a.

If, in the course of estimation of the scale parameter of the *DN* distribution a by method of quantiles, the a priori value of the shape parameter v is chosen (based on the most general considerations of physics of failure [10]) that is definitely higher than the actual value of V, the predicted average degradation rate is underestimated. On the contrary, if the a priori value of the shape parameter v is definitely lower that the actual value V, the prediction results are overestimated. And only if the chosen a priori estimation of the shape parameter is close to the actual value of the variation coefficient of the entire assembly $\hat{V}$, estimates a_i obtained by method of quantiles are around the mean estimate $\hat{a}$ with minimal dispersion and are dependence graph $a_i = f(t)$ that is as close as possible to the horizontal line around the true average.

It is recommended to average estimates a_i obtained by method of quantiles by omitting the first failures and accepting for averaging the final, most linearized section of the dependence $a_i = f(t)$, or to use the weighted average formula proposed in [1]. It must be taken into consideration that the use of statistical information on first failures causes significant errors in the estimation of the scale parameter of the *DN* distribution. No steady pattern has been identified, so estimates a_i obtained from the first failures can be both overestimated and underestimated with respect to $\hat{a}$ obtained for the entire assembly.

By using the above patterns the following method of small-sample estimation of variation coefficient of the time to failure can be formulated.

3. Method of estimation of variation coefficient based on low-level quantiles

The process of electronics degradation, along with monotone realizations (mechanical destruction in the course of thermoelectric cycling) as the result of the electric phenomena, has non-monotone realizations. Therefore, in the general case the degradation of such products is commonly considered as a process with non-monotone realizations (Figure 1). In this case the slope ratio of the average value of the determining parameters of the degradation process (inclined solid line on the graph) that occur in the product is a constant value equal to the average rate of the generalized degradation process.

$$tg\alpha = \hat{a} = \text{const}. \qquad (3)$$

The formalization of the *DN* distribution assumes that the degradation process for a set of same-type products is uniform, i.e. its average rate, mean square deviation of the rate and, subsequently, rate variation coefficient are constant (Figure 2).

The method of estimation of the variation coefficient using low-level quantiles is based on the behavior analysis of dependence graphs $a_i = f(t)$ obtained using the method of quantiles. It is considered that the best choice of the a priori value v is a choice under which the dependence graph $a_i = f(t)$ is most accurately described by a straight horizontal line, which is in compliance with the hypothesis of constant degradation rate accepted at the formalization of the *DN* distribution [4, 9] (figure 2).

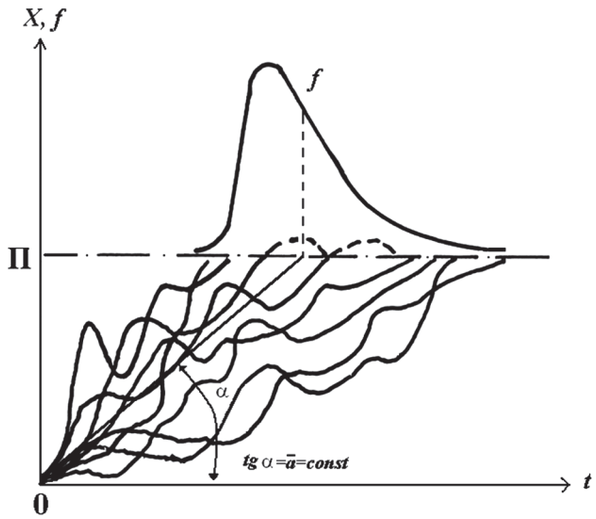


Figure 1. Graph of the *DN* distribution density formation for a product (L is the limiting value of the determining parameter that marks the onset of object failure)

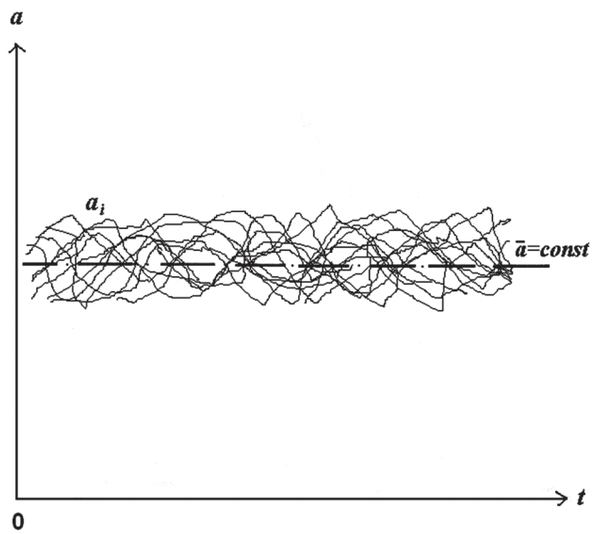


Figure 2. Graph of the theoretical dependence $a_i = f(t)$ for a set of products

In cases when the dependence graph $a_i = f(t)$ does not enable an easy conclusion regarding the best choice of a priori value v (it is especially difficult to make a choice based on the statistics of first failures), the following formal criterion can be used.

Fitting criterion of the a priori value of the shape parameter The most acceptable a priori value of the shape parameter v lies within the range of values that enable the change of the sign of the trend of the average degradation rate (h) on the graph $a_i = f(t)$.

$$h = \frac{a_n - a_1}{\bar{a}_i}, \qquad (4)$$

where a_1, a_n are the estimates of the product's degradation rate obtained based on the quantiles of the minimum and maximum levels respectively; $\bar{a}_i$ is the average value of the degradation rate estimates obtained using the method of quantiles

$$\bar{a}_i = \frac{\sum_{i=1}^{n} a_i}{n}. \qquad (5)$$

Let us illustrate the efficiency of this criterion with the example of full-scale durability tests of product samples, whose failure statistics are well described by the *DN* distribution.

An example. As an example, let us consider fatigue endurance tests of product samples made of the V-95 aluminum alloy [10]. It is required to assess the variation coefficient of the time to failure based on small samples and using the proposed method.

The first elements of the sample with the size $N = 463$ with the respective quantiles within the range from 0.0021 to 0.3131 are given in Table 1. In Table 1, the following notations are used: *r,* accumulated failure count to moment of time t_γ; t_γ, test time that corresponds to the accumulated failure count; γ, empirical failure probability.

Table 1. Data table

r	γ	t_γ, 10^3 cycle
1	0,0021	44
5	0,0107	49
10	0,0215	57
15	0,0323	59
20	0,0431	63
25	0,0539	66
30	0,0647	68
35	0,0755	73
40	0,0863	75
45	0,0971	78
50	0,1079	79
55	0,1187	82
60	0,1295	84
65	0,1403	86
70	0,1511	89
75	0,1619	91
80	0,1727	93
85	0,1835	95
90	0,1943	97
95	0,2051	99
100	0,2159	102
105	0,2267	102
110	0,2375	105
115	0,2483	106
120	0,2591	107
125	0,2699	108
130	0,2807	109
135	0,2915	111
140	0,3023	113
145	0,3131	114

Table 2. Data table

r	γ	t_γ, 10^3 cycle	$v = 0,6$		$v = 0,5$		$v = 0,4$		$v = 0,3$	
			a_i, 10^{-6} cycle^{-1}	h	a_i, 10^{-6} cycle^{-1}	h	a_i, 10^{-6} cycle^{-1}	h	a_i, 10^{-6} cycle^{-1}	h
1	0,0021	44	4,545	0,1106	5,681	0,1503	7,272	0,0358	9,318	**-0,0397**
5	0,0108	49	5,102		6,327		7,959		9,796	
10	0,0215	57	5,088		6,614		7,544		8,947	

Table 3. Data table

r	γ	t_γ, 10^3 cycle	$v = 0,6$		$v = 0,5$		$v = 0,4$		$v = 0,3$	
			a_i, 10^{-6} cycle^{-1}	h	a_i, 10^{-6} cycle^{-1}	h	a_i, 10^{-6} cycle^{-1}	h	a_i, 10^{-6} cycle^{-1}	h
15	0,0323	59	5,254	0,0092	6,44	**-0,0119**	7,797	**-0,0288**	8,915	**-0,0143**
20	0,0431	63	5,397		6,349		7,619		8,889	
25	0,0539	66	5,303		6,364		7,576		8,788	

Table 4. Data table

r	γ	t_γ, 10^3 cycle	$v = 0,6$		$v = 0,5$		$v = 0,4$		$v = 0,3$	
			a_i, 10^{-6} cycle^{-1}	h	a_i, 10^{-6} cycle^{-1}	h	a_i, 10^{-6} cycle^{-1}	h	a_i, 10^{-6} cycle^{-1}	h
15	0,0323	59	5,254	0,0148	6,44	**-0,0273**	7,797	**-0,0615**	8,915	**-0,0747**
20	0,0431	63	5,397		6,349		7,619		8,889	
25	0,0539	66	5,303		6,364		7,576		8,788	
30	0,0647	68	5,441		6,47		7,647		8,823	
35	0,0755	73	5,342		6,164		7,260		8,356	
40	0,0863	75	5,333		6,267		7,333		8,267	

Let us verify the method of variation coefficient estimation using low-level quantiles. As input data, let us take quantiles of levels from 0.0021 to 0.0215. For different values of parameter v let us define, using the method of quantiles, values a_i based on γ and r data and calculate the value of criterion h using formula (4). The values are given in Table 2.

The experimental dependence graph $a_i = f(t)$ for quantiles from 0.0021 to 0.0215 is shown in Figure 3.

Conclusions regarding parameter estimation. The change of sign of trend h occurred when $0.3 < v < 0.4$, therefore

$$v = \frac{0,3+0,4}{2} = 0,35;\ \overline{a}_i = 8,473 \cdot 10^{-6}\ \text{cycle}^{-1};$$

$$\delta_v = \frac{0,56-0,35}{0,56} = 0,375;$$

$$\delta_a = \frac{8,473 \cdot 10^{-6} - 5,9 \cdot 10^{-6}}{5,9 \cdot 10^{-6}} = 0,436.$$

The estimation error of both the shape parameter and the scale parameter are quite significant in the case of first failures. As it is known, when processing the results of dependability tests it is assumed that the first failures in a sample have the lowest weight, as their occurrence is due to serious defects not detected by final quality inspection of products. The first failures normally "fall out" of the overall statistical pattern, therefore for the purpose of further analysis we will omit them and continue the research of the effectiveness of the variation coefficient evaluation method based on quantiles of the level from 0.0323 to 0.0539. The values of a_i and h are given in Table 3.

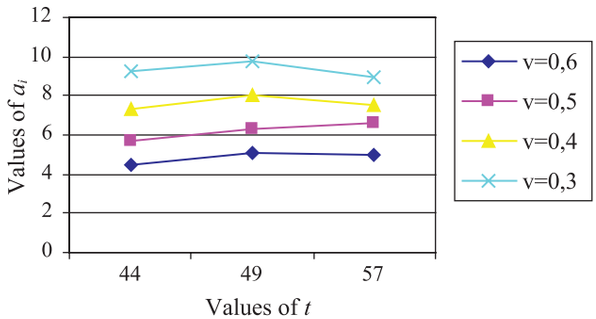


Figure 3. Experimental dependence graph $a_i = f(t)$ for quantiles from 0.0021 to 0.0215

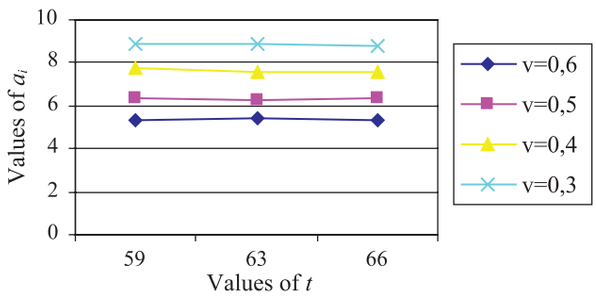


Figure 4. Experimental dependence graph $a_i = f(t)$ for quantiles from 0.0323 to 0.0539

The experimental dependence graph $a_i = f(t)$ for quantiles from 0.0323 to 0.0539 is shown in Figure 4.

Conclusions regarding parameter estimation. The last change of sign of trend h occurred when $0.5 < v < 0.6$, therefore

$$\nu = \frac{0,5+0,6}{2} = 0,55;\ \overline{a}_i = 5,851 \cdot 10^{-6}\ \text{cycle}^{-1};$$

$$\delta_\nu = \frac{0,56-0,55}{0,56} = 0,018;$$

$$\delta_a = \frac{5,9 \cdot 10^{-6} - 5,851 \cdot 10^{-6}}{5,9 \cdot 10^{-6}} = 0,008.$$

Analyzing the absolute values of trends 0.0092 and 0.0119, an additional conclusion can be made that the true value of the shape parameter v is closer to 0.6.

Let us increase the quantity of statistical information on failures to quantiles of level 0.0863. The values of a_i and h are given in Table 4.

The experimental dependence graph $a_i = f(t)$ for quantiles from 0.0323 to 0.0863 is shown in Figure 5.

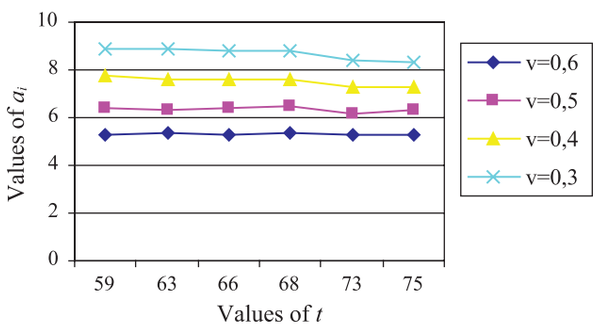


Figure 5. Experimental dependence graph $a_i = f(t)$ for quantiles from 0.0323 to 0.0864

Conclusions regarding parameter estimation. The sign of trend h did not change when the failure statistics grew and occurred again when $0.5 < v < 0.6$. Therefore

$$\nu = \frac{0,5+0,6}{2} = 0,55;\ \overline{a}_i = 5,844 \cdot 10^{-6}\ \text{cycle}^{-1};$$

$$\delta_\nu = \frac{0,56-0,55}{0,56} = 0,018;$$

$$\delta_a = \frac{5,9 \cdot 10^{-6} - 5,844 \cdot 10^{-6}}{5,9 \cdot 10^{-6}} = 0,009.$$

Under the current discreteness of variation of the a priori value of v equal to 0.1 further growth of the failure statistics does not result in more precise estimation of the shape parameter. If we assume the discreteness is equal to 0.05, the value of v could be estimated even more accurately. It must be noted that the above described process of finding the most true value of the sample estimate of the shape parameter v using the formal criterion is sufficiently algorithmic and can be successfully computerized.

Let us take a look at the graph of the average degradation rate in case the statistical information is increased to low-level quantiles of 0.3131.

In [1, 10], data are given that were obtained as the result of processing of a complete sample of V-95 products: $N = 463$, $\hat{V} = 0,56$, $\hat{S} = 169 \cdot 10^3$ cycle, $\hat{a} = 5,9 \cdot 10^{-6}$ cycle^{-1}.

The estimates of the variation coefficient of the time to failure per low-level quantiles using the proposed formal criterion are very close ($v = 0.55$; $\overline{a}_i = 5.844 \cdot 10^{-6}$ cycle^{-1}) to the estimates obtained experimentally using a complete sample.

Figure 6 shows the dependence graph of the DN distribution scale parameter estimate obtained per quantiles from 0.0021 to 0.3131 for $v = 0.57$.

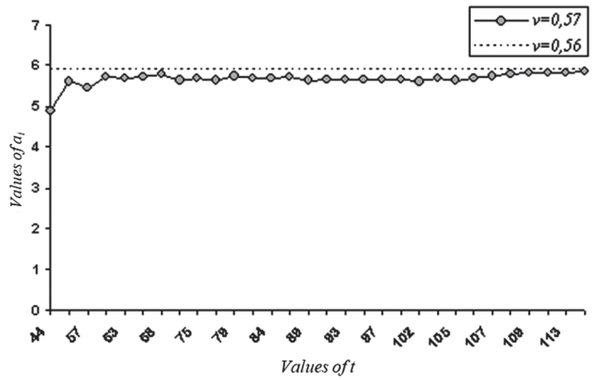


Figure 6. Experimental dependence graph $a_i = f(t)$ for quantiles from 0.0021 to 0.3131

As it can be seen, while the a priori value of the shape parameter $v = 0.57$ is almost completely identical to the sample estimate of the variation coefficient $\hat{V} = 0,56$, the dependence graph $a_i = f(t)$ is a sufficiently straight line slightly below estimate $\hat{a} = 5,9 \cdot 10^{-6}$ cycle^{-1} obtained per the complete sample. As it was expected, the exception is the first failures that underestimate the average degradation rate and do not match the general trend.

4. Conclusions

The proposed method of quantile-based estimation of the variation coefficient of the time to failure enables – in the context of limited failure statistics – using ultralow level quantiles for sufficiently accurate identification of not only the variation coefficient of the time to failure and *DN* distribution parameters, but also make conclusions regarding the feasibility and legitimacy of equalization (description) of the considered sample using this diffusion distribution, i.e. it can be used as a kind of criterion of compliance of the empirical failure distribution under consideration with the chosen theoretical dependability model. The above described process of finding the most true values of the variation coefficient of the time to failure using the formal criterion can be computerized.

References

[1]. Strelnikov VP, Fedukhin AV. Otsenka i prognozirovanie nadezhnosti elektronnykh elementov i sistem [Estimation and prediction of the dependability of electronic elements and systems]. Kiev: Logos; 2002 [in Russian].

[2]. Savchuk VP. Bayesovskie metody statisticheskogo otsenivaniya nadezhnosti tekhnicheskikh obektov [Bayesian methods of statistical evaluation of technical objects dependability]. Moscow: Nauka; 1989 [in Russian].

[3]. Prokhorenko VD, Golikov VF. Uchet apriornoy informatsii pri otsenke nadezhnosti [Accounting for a priori information in dependability estimation]. Moscow: Nauka i teknika; 1978 [in Russian].

[4]. GOST 27.201-81. Reliability in the equipment. Assessment of indicators of reliability at small number of observations with the use of additional information. Moscow: Izdatelstvo standartov; 1981 [in Russian].

[5]. Strelnikov VP. Prognozirovanie resursa izdeliy elektronnoy tekhniki [Predicting of electronic products' lifetime]. Dependability 2014;4(12):43-48 [in Russian].

[6]. Strelnikov VP. Metodicheskie pogreshnosti rascheta nadezhnosti sistem [Systematic errors of system dependability calculation]. Dependability 2005;3(12):41-46 [in Russian].

[7]. Strelnikov VP. Metodicheskie pogreshnosti rascheta nadezhnosti elektronnykh elementov i sistem [Systematic errors of electronic components and systems dependability estimation]. Dependability 2009;2:27-32 [in Russian].

[8]. Strelnikov VP. Zakonomernosti izmeneniya narabotki mezhdu otkazami tekhnicheskikh sistem v protsesse ekspluatatsii [Change trends of time between failures in technical systems in operation]. Dependability 2011;1(36):17-22 [in Russian].

[9]. Pogrebinsky SB, Strelnikov VP. Proektirovanie i nadezhnost mnogoprotsessornykh EVM [Design and dependability of multiprocessor computers]. Moscow: Radio i svyaz; 1988 [in Russian].

[10]. GOST 27.005-97. Industrial product dependability. Failure models. General concepts. Kiev: Gosstandart Ukrainy; 1997 [in Russian].

About the authors

Fedukhin Alexander Viktorovich, Head of Laboratory of dependable computer systems for critical technologies and infrastructures, Institute of Mathematical Machines and Systems Problems, NAS of Ukraine, Doctor of Engineering, Senior Researcher, phone: +380679898306, avfedukhin@gmail.com

Cespedes Garcia Natalia Vasilievna, Bench Scientists, Laboratory of dependable computer systems for critical technologies and infrastructures, Institute of Mathematical Machines and Systems Problems, NAS of Ukraine, phone: +380932568725, nata05805@gmail.com

Received on: 03.07.2018

Set of indicators for dependability evaluation of gas compression units

Igor R. Baikov, *Ufa State Petroleum Technological University, Ufa, Russia*
Sergey V. Kitaev, *Ufa State Petroleum Technological University, Ufa, Russia*
Olga V. Smorodova, *Ufa State Petroleum Technological University, Ufa, Russia*


Igor R. Baikov


Sergey V. Kitaev


Olga V. Smorodova

Abstract. *The paper is dedicated to the improvement of the evaluation methods of one of the most important operating characteristics of gas compression units (GCUs), i.e. dependability, under the conditions of decreasing pipeline utilization rate. Currently, the dependability of units is characterized by a set of parameters based on the identification of the time spent by a unit in certain operational state. The paper presents the primary findings regarding the dependability coefficients of GPA-Ts-18 units, 41 of which are operated in multi-yard compressor stations (CSs) of one of Gazprom's subsidiaries. The dependability indicators (technical state coefficient, availability coefficient, operational availability coefficient) identified as part of the research are given as well. GCUs were classified into groups depending on the coefficient values. The feasibility of using integral indicators in the analysis of GCU groups' dependability was examined. It was proposed to use confidence intervals for identification of the integral level of dependability of the operated GCU stock and the ways of maintaining the operability of units under the conditions of decreasing main gas pipeline utilization rate. The Gini index was suggested for the purpose of generalized estimation of GCU groups' dependability. It is shown that the advantage of the Gini coefficient is that is allows taking into account the ranks of the analyzed features in groups. The graphic interpretation of the findings was executed with a Lorenz curve. The paper implements the sigma rule that characterizes the probability of the actual coefficient value being within the confidence interval, i.e. prediction limits (upper and lower), within which the actual values will fall with a given probability. The confidence intervals were identified by the type of coefficients distribution and a standard deviation, . A histogram of an interval range of technical utilization coefficient distribution is given as an example. Testing of the hypothesis of the distribution type at confidence level 0.95 showed that the distribution of coefficients is normal. Using the moment method, the mathematical expectation and mean square deviation for the distribution of the values of each type of dependability indicators were established. Using the sigma rule, all extreme outliers among the GCUs in terms of the level of factor attribute were excluded from the body of input data. All units whose factor attribute value does not fall in the interval were excluded. According to the three sigma rule, 3 and 2 GCUs did not fall in the confidence interval ($\mu \pm 3\sigma$) in terms of the utilization factor and availability factor respectively. The performed analysis of causes of low availability coefficients of the above GCUs showed that the systems had been long in maintenance. The paper sets forth summary data on the maximum allowable value of the Gini index of dependability coefficients (C_{TU}, C_A, C_{OA}) depending on the sample size (the complete sample of 41 units and samples with the interval of 1, 2, 3 sigma). In case of higher values of Gini index it is recommended to adopt measures to individual units in order to improve the dependability of the operated GCU stock.*

Keywords: *gas compression unit, dependability, failure, indicator, operability, three sigma rule.*

For citation: *Kitaev SV, Baikov IR, Smorodova OV. Set of indicators for dependability evaluation of gas compression units. Dependability 2018;4: 16-21. DOI: 10.21683/1729-2646-2018-18-4-16-21*

Introduction

One of the most important characteristics of gas compression units (GCUs) is dependability. Dependability of GCU as a whole is defined by the dependability of its elements, support systems and the nature of their interaction [1-3].

The paper sets forth a study of a set of dependability coefficients and the development of indicators to differentiate GCUs by the indicators' value. The developed coefficients may be useful for studying the performance of GCUs operated in multi-yard compressor stations (CSs).

Research scope analysis

The choice of the research object is based on the need to ensure the operability of a gas compression system in emergency mode. It is known that the transfer of mains gas by units of medium and small single capacity improves the flexibility of a system with guaranteed redundancy. At the same time, an emergency shutdown of one of the units causes minimal harm to the process.

The emergency shutdown of a large unit can cause much greater negative effects. There are 5 standard sizes of units of high single capacity (Table 1) operated by Gazprom's subsidiaries.

More than 77% of all 79 units are operated by Gazprom Transgaz Yugorsk.

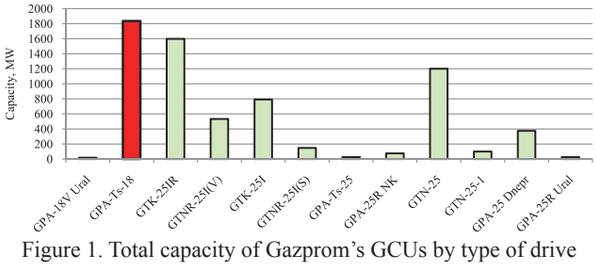


Figure 1. Total capacity of Gazprom's GCUs by type of drive

In accordance with the identified structure of the GCU stock, the GCU-Ts-18, 41 of which are operated in multi-yard (9 compressor yards) CSs, were chosen as the research object. This is a conversion GCU with an aircraft gas turbine. At the time of the research the total operating time of the units was from 20 thousand to 136 thousand hours (113 thousand hours on the average).

Defining GCU dependability coefficients

Currently, the gas turbine (GT) dependability is evaluated using a system of indicators [4, 5] that are based on the identification of the time the unit is in a particular operational state: total operation time T_o during the reporting period T_c; time of the unit on stand-by T_{sb}; time of the unit being under scheduled repair T_{sr}; GCU downtime T_d during the reporting period T_c:

- technical utilization coefficient, C_{TU}
- availability coefficient, C_A
- operational availability coefficient, C_{OA}
- mean time between failures during the reporting period, T_F
- restoration time coefficient, C_R.

Many authors [1, 2] demonstrate that the failure rate that defines the dependability of equipment operation is primarily associated with the GCUs' aging process. Meanwhile, preventive and diagnostic maintenance measures can not only help sustain the GCUs' technical condition, but also correct it. The time mode of the GCU being in each operational state is not directly connected to the total operating time of the unit and is an additional indicator for identifying its dependability. Their mutual independence is confirmed by the value of the cross-correlation coefficient: it lies in the range (-0.094; 0.126) for various coefficients, which confirms the absence of a significant correlation.

Figure 2 shows the distribution of dependability indicators' values for GPA-Ts-18. The analysis is based on the results of 2 years of units' operation.

Figure 3 shows structure diagrams of dependability coefficients' distribution by intervals. Interval estimation of

Table 1. Primary information on gas compression units of high single capacity, Gazprom

No.	Type of GCU	Number of GCUs	Single capacity	Proportion of total capacity in a group of large units
			MW	%
1	GPA-18V Ural	1	18	0.3
	GPA-Ts-18	102	18	27.3
2	GTK-25IR	72	22.2	23.8
	GTNR-25I(V)	24	22.2	7.9
3	GTK-25I	33	24	11.8
4	GTNR-25I(S)	6	24.5	2.2
5	GPA-Ts-25	1	25	0.4
	GPA-25R NK	3	25	1.1
	GTN-25	48	25	17.8
	GTN-25-1	4	25	1.5
	GPA-25 Dnepr	15	25	5.6
	GPA-25R Ural	1	25	0.4
	TOTAL	310		

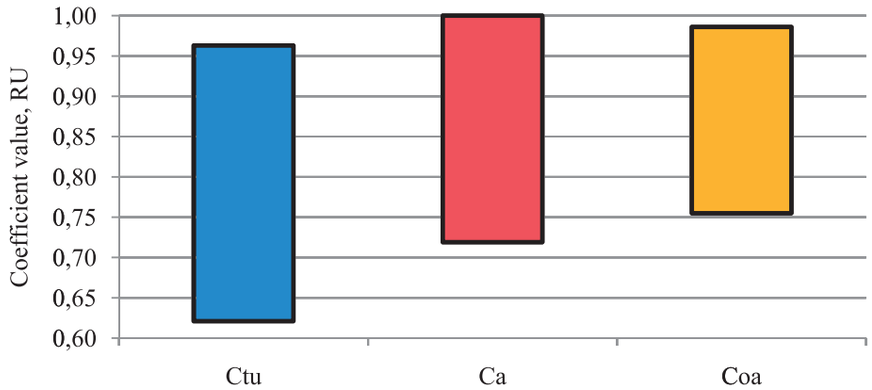


Figure 2. Distribution of dependability indicators values for GCU

dependability coefficients shown in the diagrams was used to develop and justify differentiation indicators.

Results show that the technical utilization coefficient for GPA-Ts-18 units is 0.621÷0.963; availability coefficient is 0.719÷1.0; operational availability coefficient is 0.755÷0.986. Mean time between failures is $T_F = 2900$ hours; mean value of restoration time coefficient is $C_R = 70$ hours.

The value of technical utilization coefficient for converted GCUs should be no less than 0.94, availability coefficient should be no less than 0.98, mean time between failures should be no less than 3500 hours [6].

Thus, most of the GCUs from the examined group have coefficients lower than the values established by GOST [6]: 93% of units in terms of technical utilization coefficient, 10% of units in terms of availability coefficient, 95% of units in terms of operational availability coefficient, 76% of units in terms of mean time between failures.

Definition of GCU integral dependability indicators

The deviation of the coefficients from the standard values is due to a lower main gas pipeline utilization rate that is below the design value and gas being transferred by fewer GCUs. In such complicated conditions, there is a need for additional methods to identify the level of dependability of the GCU operational stock.

In order to make a decision on the general dependability of the enterprise's GCUs for repair planning and operating modes optimization, an integrated assessment of the dependability indicators of the gas compression equipment stock as a whole was performed. The Gini index was used as an integral indicator.

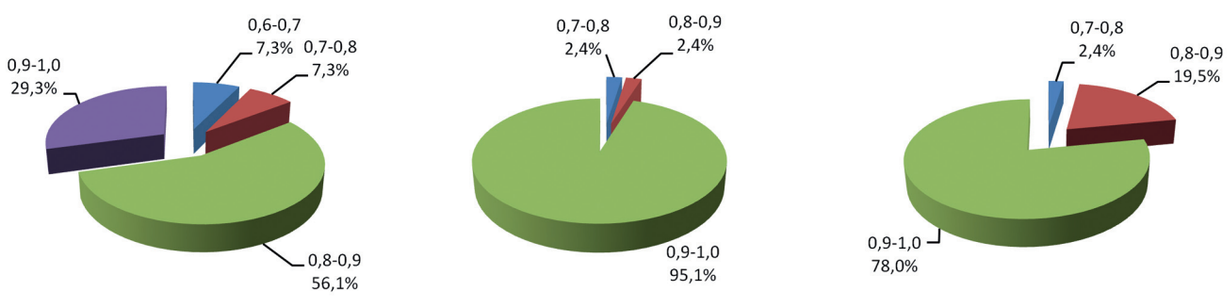
Initially, the Gini coefficient was introduced in the economic science as a measure of the population's income concentration [7] to evaluate the degree of inequality between certain social groups. The indicator was used by the authors of [8, 9] in the oil and gas industry for differentiating equipment according to technical and operating conditions in the extraction, pipeline transportation and processing of hydrocarbons. The Gini coefficient can theoretically range from 0 to 1. The closer the value is to one, the greater the differentiation of equipment by the studied indicator is.

Regarding the assessment of GCUs' dependability level differences, the Gini coefficient will show the differentiation of GCUs by dependability level that is defined by the coefficients of technical utilization, availability and operational availability. The Gini coefficient is calculated using the formula (Figure 4a):

$$K_L = 1 - 2\sum_i X_i \mathrm{cum}\, Y_i + \sum_i X_i Y_i,$$

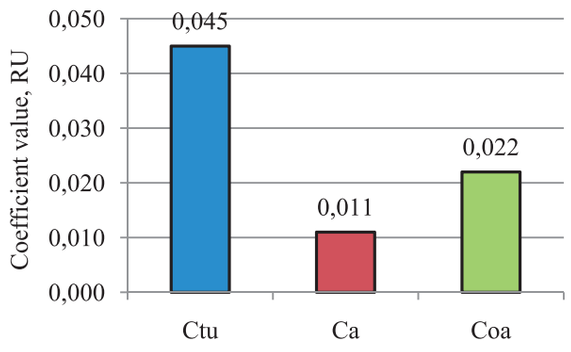
where X_i is the proportion of the GCUs in group i; Y_i is the proportion of the group i in the overall level of coefficients; $cum\, Y_i$ is the cumulative (calculated as progressive total) proportion of the coefficients.

The Gini coefficient is calculated based on the data on GCUs' distribution by the level of dependability indicators. The entire set is divided into N groups with an equal number of GCUs, and the proportion of each group in the total sum of

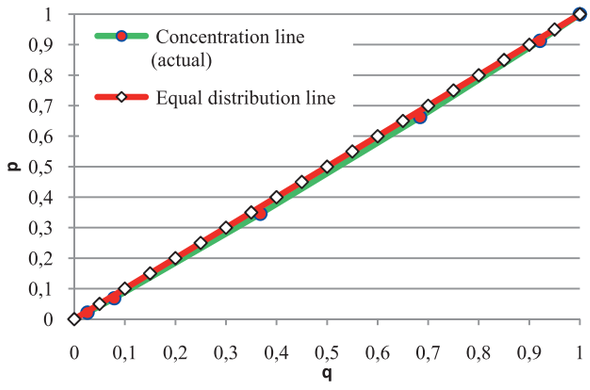


a) technical utilization coefficient b) availability coefficient c) operational availability coefficient

Figure 3. Interval structure of GCU dependability coefficients



a – values of concentration coefficients



b – Lorenz curve of distribution of technical utilization coefficient

Figure 4. Differentiation index for GCU sample (41 units)

the coefficients is identified. A concentration curve (Lorenz curve) was constructed based on the cumulative specific weights (frequencies) by the number of GCUs and specific weights in the total sum of indicators.

The cumulative portion of the groups in the total sum of the indicator (from 0% to 100% or from 0 to 1) is represented by the vertical axis. The cumulative portion of the GCU groups in the total amount (from 0% to 100% or from 0 to 1 as well) is represented by the horizontal axis. If the indicator was distributed equally, each group of GCUs would have exactly the same part of the total sum of the indicator as its percentage. On the graph, this is depicted by the diagonal line called the line of equal distribution.

The actual indicator distribution is the concave concentration line below the diagonal. The further this line is from the diagonal, the more unequal is the distribution of the indicator (the higher the level of concentration). Graphs of technical utilization (Figure 4b), availability and operational availability coefficients' values concentration curves were constructed based on the results of the calculation.

In theory, the characteristic of concentration of coefficients' values may coincide with the line of equal distribution, in which case the differentiation index (Gini) will be equal to zero, and the level of GCUs' dependability in the group will be equal.

As Figure 4a shows, the calculated values of differentiation indexes (Gini) are rather low, which indicates that the difference in the GCUs' level of dependability is insignificant. However, regarding the major part of GCUs, the values of the coefficients are lower than the standard values. At the same time, an uncertainty remains regarding the decision on the choice of strategy of GCU stock operating method in conditions of low main gas pipeline utilization rate.

Statistical analysis of GCU dependability level

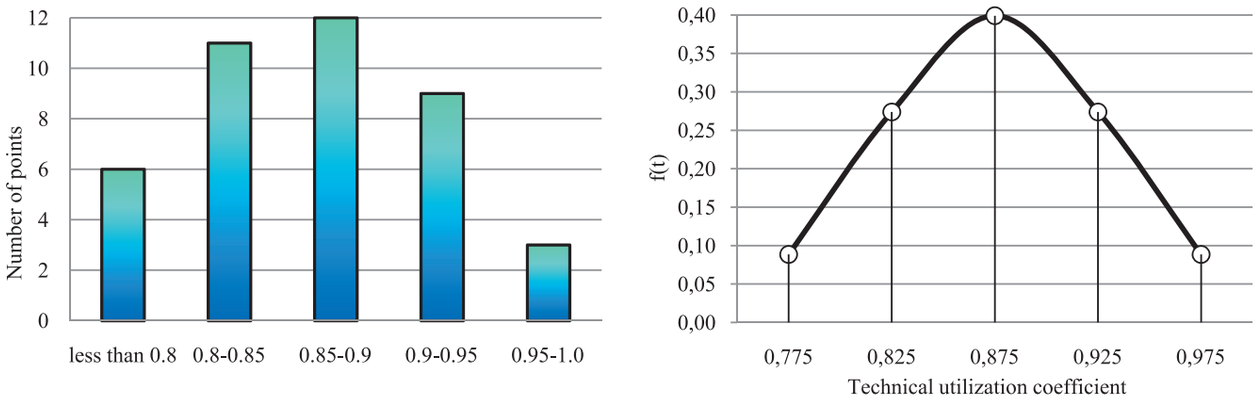
In the dependability theory, the sigma rule characterizes the probability of the next actual value being within the confidence interval. The confidence interval helps identify areas that should be addressed to change the trend and make an informed decision (for example, determine the strategy for GCU repair and maintenance). In regard to dependability coefficients, the confidence interval is the prediction limits (upper and lower), within which with a given probability the actual values will lie.

When the confidence interval is:

- 3 sigma, than there is a 0.3% probability that the value of the parameter lies outside the confidence interval;
- 2 sigma, than there is a 4.5% probability that the value of the parameter lies outside the confidence interval;

Table 2. Evaluation of the body of data of coefficients CTU, CA, COA according to the sigma rule

Designation of coefficient	Intervals for body of coefficient data	Intervals of factor indicator values	Number of units in an interval	Specific weight of units in an interval in the total number, %
1	*2*	*3*	*4*	*5*
C_{TU}	$\bar{y}-\sigma \le y_i \le \bar{y}+\sigma$	$0.817 \le y_i \le 0.933$	30	73.2
	$\bar{y}-2\sigma \le y_i \le \bar{y}+2\sigma$	$0.760 \le y_i \le 0.990$	37	90.2
	$\bar{y}-3\sigma \le y_i \le \bar{y}+3\sigma$	$0.702 \le y_i \le 1.048$	38	92.7
C_A	$\bar{y}-\sigma \le y_i \le \bar{y}+\sigma$	$0.977 \le y_i \le 0.993$	12	29.3
	$\bar{y}-2\sigma \le y_i \le 2\bar{y}+2\sigma$	$0.970 \le y_i \le 1.0$	39	95.1
	$\bar{y}-3\sigma \le y_i \le \bar{y}+3\sigma$	$0.962 \le y_i \le 1.008$	39	95.1
C_{OA}	$\bar{y}-\sigma \le y_i \le \bar{y}+\sigma$	$0.889 \le y_i \le 0.962$	28	68.3
	$\bar{y}-2\sigma \le y_i \le 2\bar{y}+\sigma$	$0.852 \le y_i \le 0.998$	39	95.1
	$\bar{y}-3\sigma \le y_i \le \bar{y}+3\sigma$	$0.716 \le y_i \le 1.035$	41	100



a) histogram of distribution b) normal density

Figure 5. Distribution of technical utilization coefficient

- 1 sigma, than there is a 31.7% probability that the value of the parameter lies outside the confidence interval.

The confidence interval is constructed using the distribution of technical utilization coefficient and a standard deviation, σ. Figure 5a shows a histogram of an interval range of technical utilization coefficient distribution.

Tests of the hypothesis of the distribution type at confidence level 0.95 showed that the distribution of coefficients is normal (Figure 5). Using the moment method, the mathematical expectation and mean square deviation for the intervals of coefficients' values were calculated (Figure 5b).

Using the sigma rule, all extreme outliers among the GCUs in terms of the factor attribute level were excluded from the input data (Table 2).

Table 3 shows calculated values of the differentiation index (Gini) for samples with intervals 1σ, 2σ, 3σ.

In accordance with the Chebyshev's theorem, the "three sigma" rule is widely used in engineering to ensure the dependability of equipment stock operation with a sufficient degree of probability. Not all values fall in the interval (μ±3σ). Three GCUs can be excluded having the lowest technical utilization coefficient values of 0.621 (GPA no. 3 KTs-10), 0.663 (GPA no. 1 KTs-2) and 0.681 (GPA no. 4 KTs-7).

Table 3. Calculated values of the differentiation index (Gini) for samples with intervals 1σ, 2σ, 3σ

Interval of values	Gini index		
	for C_{TU}	for C_A	for C_{OA}
1 sigma	0.016	0	0.006
2 sigma	0.029	0	0.016
3 sigma	0.032	0	0.019

According to the three sigma rule, values of availability coefficient 0.816 (GPA no. 1 KTs-2) and 0.719 (GPA no. 4 KTs-7) did not fall in the confidence interval.

The performed analysis of causes of low values of availability coefficient of the above GCUs showed that the units had been long in maintenance due to failure and delay in the delivery of spare parts.

Figure 6 shows summary data on the Gini differentiation index of dependability coefficients (C_{TU}, C_A, C_{OA}) depending on the sample size (the complete sample of 41 units and samples with the interval of 1σ, 2σ and 3σ).

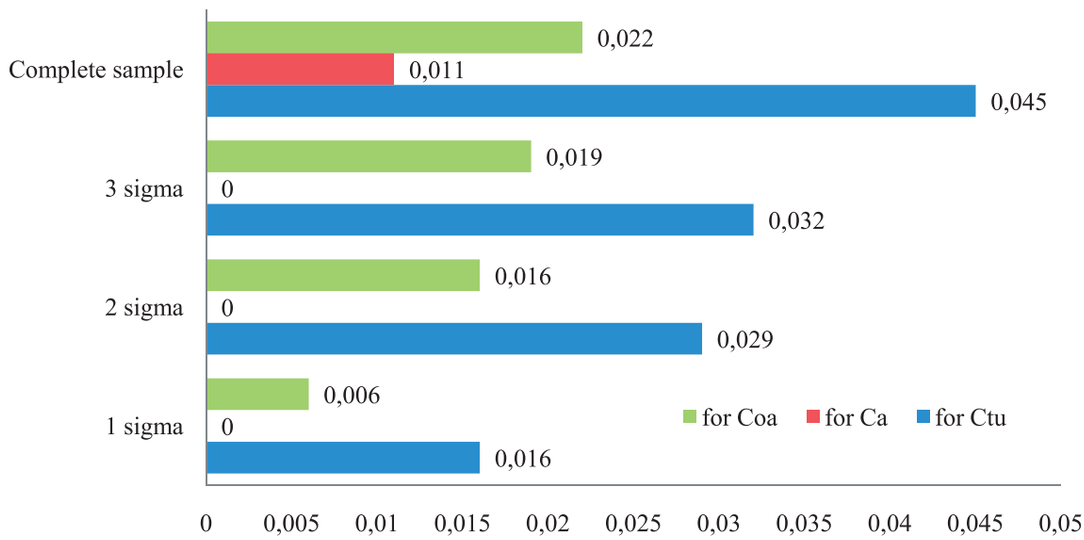


Figure 6. Histogram of Gini differentiation index distribution

To sum up, according to the "three sigma" rule, the value of the Gini index for converted GPA-Ts-18 units should be no greater than 0.032 for the technical utilization coefficient, no greater than 0 for the availability coefficient and no greater than 0.019 for the operational availability coefficient.

The advantage of the Gini coefficient over the arithmetic mean value of the coefficients for the analyzed groups is that the indicators are calculated more accurately. The Gini index allows taking into account the ranks of the analyzed attributes in groups (eliminating the influence of isolated GCUs) and identifying the degree of differentiation of GCU groups in terms of dependability.

Conclusion

1. The values of dependability coefficients were determined (a sample of 41 GCU-Ts-18 units was analyzed): the technical utilization coefficient is 0.621÷0.963; the availability coefficient is 0.719÷1.0; the operational availability coefficient is 0.755÷0.986. Meanwhile, most of the GCUs from the examined group have coefficients lower than the values established by GOST, the reason being the decrease in main gas pipeline utilization rate.

2. Indicators of differentiation of GCU groups by dependability level (Gini) for technical utilization coefficient, availability coefficient and operational availability coefficient were suggested. The advantage of Gini coefficient is that is allows taking into account the ranks of the analyzed attributes in groups making the calculations of differentiation level more accurate.

3. According to the "three sigma" rule, the value of Gini index for GPA-Ts-18 units with the total operating time being up to 136 thousand hours should be no greater than 0.032 for the technical utilization coefficient, no greater than 0 for the availability coefficient and no greater than 0.019 for the operational availability coefficient. In case of higher values of Gini index it is recommended to adopt measures to individual units in order to improve the dependability of the operated GCU stock.

References

[1]. Baikov IR, Smorodova OV, Kitaev SV. The pipeline gas turbine set reliability parameters estimation. Neftegazovoe delo 2017;1:95-107, <http://ogbus.ru/issues/1_2017/ogbus_1_2017_p95-107_BaikovIR_ru.pdf> [in Russian].

[2]. Kitaev SV, Kuznetsova MI. Statisticheskoe modelirovanie pokazateley nadezhnosti gazoturbinnykh ustanovok metodom «Monte-Karlo» [Monte Carlo modeling of gas turbine sets dependability indicators]. Gazovaya promyshlennost 2014;5:101-103 [in Russian].

[3]. Smorodov EA, Kitaev SV. Metody rascheta koeffitsientov tekhnicheskogo sostoyaniya GPA [Methods of calculating the coefficients of technical state of GCUs. Gazovaya promyshlennost 2000;S:29-31 [in Russian].

[4]. GOST 27.002-2015. Dependability in technics. Terms and definitions. Moscow: Standartinform; 2015 [in Russian].

[5]. GOST R 52527-2007 (ISO 3977-9:1999). Gas turbine plants. Reliability, availability, maintainability and safety. Moscow: Standartinform; 2006 [in Russian].

[6]. GOST R 54404. Gas pumping units driven with gas turbine. General specifications. Moscow: Standartinform; 2012 [in Russian].

[7]. Kharchenko LP, Ionin VG, Glinsky VV et al. Statistika [Statistics]. Moscow: Infa-M; 2008 [in Russian].

[8]. Baikov IR, Kitaev SV, Valiev AN, Zuev AS, Starostin VV. Energosberezhenie pri ekspluatatsii fonda tsentrobezhnykh elektronasosov na neftyanykh promyslakh [Energy saving in operation of centrifugal electric pumps in oil fields]. Transport i khranenie nefteproduktov i uglerodnogo syr'ya. 2011;4:23-26 [in Russian].

[9]. Smorodova OV, Kitaev SV, Sergeeva KV. The technological oil refining units ranging on the base of industrial safety generalized criteria. Neftegazovoe delo 2017;4:165-179, <http://ogbus.ru/issues/4_2017/ogbus_4_2017_p165-179_SmorodovaOV_ru.pdf> [in Russian].

About the authors

Igor R. Baikov, Doctor of Engineering, Professor, Head of the Department of Industrial Thermal Power Engineering, Faculty of Pipeline Transport, Ufa State Petroleum Technological University, Ufa, Russia, e-mail: pte@rusoil.net

Sergey V. Kitaev, Doctor of Engineering, Senior Lecturer, Professor of the Department of Transport and Storage of Oil and Gas, Faculty of Pipeline Transport, Ufa State Petroleum Technological University, Ufa, Russia, e-mail: svkitaev@mail.ru

Olga V. Smorodova, Candidate of Engineering, Associate Professor, Senior Lecturer of the Department of Industrial Thermal Power Engineering, Faculty of Pipeline Transport, Ufa State Petroleum Technological University, Ufa, Russia, e-mail: olga_smorodova@mail.ru

Received on 30.03.2018

Original article
Dependability, vol. 18 no.4, 2018 DOI: 10.21683/1729-2646-2018-18-4-22-27

Research of operational dependability of automotive engines

Yuri V. Bazhenov, *AG and NG Stoletov Vladimir State University, Vladimir, Russia*
Mikhail Yu. Bazhenov, *AG and NG Stoletov Vladimir State University, Vladimir, Russia*

Yuri V. Bazhenov

Mikhail Yu. Bazhenov

Abstract. *The problem of increasing the dependability of the engine, which is the most complex and expensive unit of an automotive vehicle, cannot be solved without objective and reliable information on the failures and malfunctions of its components, their causes, actual life, as well as the factors affecting such indicators in real operational conditions. Manufacturing factories do not always have such information, hence design deficiency failures associated with design and development flaws are among the most common causes of loss of engine operability.* **The aim** *of this paper is to study the engine operational dependability using the results of their maintenance and repair.* **The methods** *are based on operational tests of engines that yield the most complete and objective information on their dependability, as they were conducted in typical operational conditions of automobile operating companies in the course of vehicle maintenance and repair. The results of the studies processed with the standard Statistica 6.0 are represented in the form the statistical evaluations of the dependability of primary structural engine components (times to failure, changes in the probability of no-failure depending on the travelled distance). The analysis of the obtained information allows estimating the level of actual dependability of the engine, identifying design flaws, developing specific measures aiming to increase operational dependability. Information obtained during such tests is useful not only to the engine manufacturers, but to the operators as well, as it enables a scientific substantiation of the norms of operability. For the purpose of identification and localization in the process of maintenance and repair of specific engine malfunctions, the paper substantiates a set of diagnostic parameters and their standard values.* **Conclusions.** *The research allowed elaborating a set of diagnostic parameters for evaluation of the technical condition of primary engine systems (cylinder-piston group, crank and gas distributing mechanisms) that define and limit its dependability. The application of the findings in the automobile maintenance and repair processes enables a significant improvement of the engines' operational dependability and reduction of the costs of ensuring their operability.*

Keywords: *engine, automobile, dependability, failure, operation time, structural parameter, diagnostic parameter, technical condition.*

For citation: *Bazhenov YuV, Bazhenov MYu. Research of operational dependability of automotive engines. Dependability 2018;4: 22-27. DOI: 10.21683/1729-2646-2018-18-4-22-27*

To solve the problem of ensuring a high dependability level and operability of technical systems, different types of information on their operation conditions, acting loads, mode and causes of failures and malfunctions are required. The availability of such information is a prerequisite of the improvement of system dependability at all lifecycle stages and the basis for the development of measures to improve the design, processes of its manufacture and operation. This all applies to the internal-combustion engine that is the most complex and expensive unit of a vehicle, which account for up to 20% of all its failures.

Engines manufacturing factories do not always have reliable information on malfunctions arising during operation, causes of failures, operation time to limit state and other indicators characterizing operating dependability of their products. As a result, in actual operating conditions, among causes of engine failures, there are design deficiency failures caused by the imperfection of their design and engineering.

The tests (development, research, acceptance, validation, etc.) are a source of reliable information on engines dependability, as well as any other mechanism and system

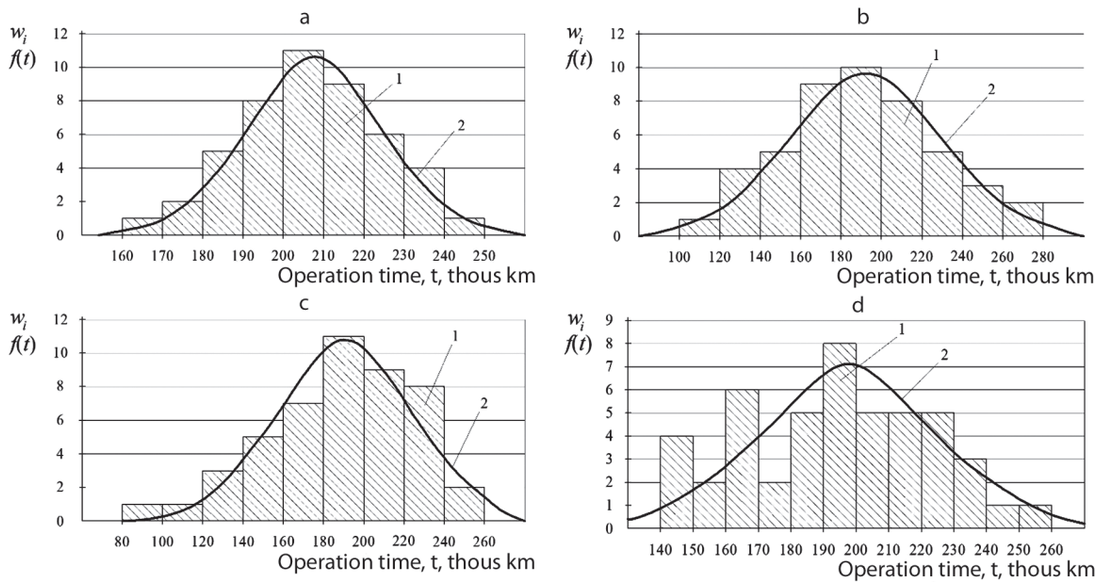


Figure 1. Histograms (1) and theoretical curves (2) of times to failure distribution: a) engine block; b) cylinder head; c) piston block; d) cranked shaft

Table 1. Statistical estimates of the numerical characteristics of the engines dependability

no.	Names of structural engine components	Mean liftime t_{mlt}, thous. km	Mean square deviation, y, thous. km	Variation coefficient, н
1	Engine block	203.7	34.5	0.169
2	Cranked shaft	198.4	39.2	0.198
3	Connection shaft	141.8	41.7	0.294
4	Distributive shaft	194.6	26.8	0.138
5	Piston block	191.6	35.3	0.184
6	Piston rings	148.2	45.2	0.304
7	Bottom-end bearing	164.0	41.4	0.254
8	Cranked shaft bearing	166.0	40.6	0.245
9	Connecting rod bush	195.9	52.9	0.270
10	Piston pin	186.2	37.2	0.200
11	Valve guide	154.6	34.0	0.220
12	Deflation valve	169.0	34.8	0.206
13	Hydraulic tappet	131.8	39.8	0.302
14	Cylinder head	193.6	39.0	0.201

of the vehicle. The most objective and exhaustive information on the engines dependability is produced by operation tests that are carried out in typical vehicle operating conditions. Information obtained during such tests is useful not only for the engine manufacturers, but for the operators as well, as it enables a scientific substantiation of the norms of operability.

As part of this paper, a research on engines dependability was carried out in actual operating conditions with registration of condition data during technical maintenance and repair of vehicles. The ZMZ-4063.10 engine produced by the Zavolzhsky engine plant and installed on the GAZelle family vehicles was taken as the object of research. A large amount of information on malfunctions and engines failures arising during vehicles operation, has been collected.

The findings on the operation dependability of primary structural engine components, processed with the standard program Statistica 6.0, are presented in Table 1, and partially in the forms of histograms and theoretical curves of times to failure distribution in Figure 1.

The curves types as well as the calculated values of variation coefficient н show that distribution of times to failure of the engine components is described by the normal law. Test of the hypothesis on experimental data belonging to the normal probability with the Pearson fitting criterion χ^2 confirmed its validity.

One of the main indicators that evaluates the dependability of structural engine components is the probability of its fail-safe operation $P(t)$ or failure $F(t)$ within the limits of operating time. Table 2 shows the processing results of statistical data on dependability of primary components of the studied engines, which clearly show the change of probability of their failures in operating time t.

Analyzing the data presented in Table 2, some conclusions can be made regarding the operation dependability of components of the studied engines. The probability of fail-safe operation of components for the initial intervals of operating time from 0 to 90 thous. km is at a sufficiently high level. During this period of operation there is a low probability of failure of the deflation valves and hydraulic tappets of gas distribution mechanism which are exposed to high mechanical and thermal loads during engine operation. The probability of fail-safe operation of connection shaft, cylinder head gasket, piston rings, bottom-end bearing and valve guide decreases significantly by the operating time of 154 thous. km. Within the interval of operation time from 154 to 218 thous. km, there is a sharp increase in the engine failure probability which for different components ranges from $F(t) = 0.620$ (connecting rod bush) to $F(t) = 0.995$ (deflation valve). Practically all engine components exhaust their lifespan by the operating time of 250 thous. km. By this operating time the failure probability of the base component, the engine block, reaches $F(t) = 0.911$, which indicates the requirement for an overhaul or decommissioning.

Among the reasons for this level of the engine operation dependability in addition to the design and manufacture factors, the effects of the operating conditions should be noted: road condition, storage, environmental conditions, infrastructure and others. The operating conditions include the maintenance system with control and diagnostic, preventive and repair measures aimed at ensuring engine efficiency.

To ensure reliable engine operation and reduce the cost of maintenance operations after failures, most of them must be prevented as part of scheduled maintenance. Therefore, during maintenance, it is required to have the information on the engine technical condition, on hidden and imminent

Table 2. Probabilities of failure of primary engine components ZMZ-4063 in operating time

no.	Component title	Probability of failure $F(t)$ in time, thous. km						
		58	90	122	154	186	218	250
1	Engine block	0	0.001	0.009	0.075	0.304	0.661	0.911
2	Cranked shaft	0	0	0.004	0.065	0.336	0.749	0.961
3	Connection shaft	0.022	0.107	0.317	0.615	0.856	0.966	0.995
4	Distributive shaft	0	0	0.003	0.065	0.375	0.810	0.981
5	Piston block	0	0.002	0.024	0.143	0.437	0.773	0.951
6	Piston rings	0.023	0.099	0.281	0.551	0.799	0.939	0.988
7	Bottom-end bearing	0.005	0.037	0.161	0.405	0.702	0.904	0.981
8	Cranked shaft bearing	0.004	0.030	0.139	0.384	0.689	0.901	0.981
9	Connecting rod bush	0.005	0.023	0.081	0.214	0.426	0.620	0.847
10	Piston pin	0	0.005	0.042	0.194	0.438	0.804	0.957
11	Valve guide	0.002	0.031	0.169	0.493	0.807	0.969	0.997
12	Deflation valve	0.053	0.221	0.5314	0.827	0.962	0.995	0.999
13	Hydraulic tappet	0.071	0.229	0.493	0.759	0.923	0.984	0.998
14	Cylinder head	0	0.001	0.033	0.155	0.423	0.735	0.926
15	Cylinder head gasket	0.006	0.069	0.332	0.729	0.951	0.996	0.999

Table 3. Structural and evaluating diagnostic parameters of the ZMZ-4063 engine

no.	Diagnostic parameter	Structural parameter
1	Pressure at the end of a compression stroke, S_1	• Clearance between the ring and the 1-st compression ring by groove width, Y_1 • Clearance in gap of the 1-st compression ring, Y_2 • Clearance between the ring and the 2-nd compression ring by groove width, Y_3 • Clearance in gap of the 2-nd compression ring, Y_4 • Clearance between the piston and the engine block, Y_5 • Valve plug-to-guide bush clearance, Y_6 • Valve plug-to-guide bush clearance Y_7
2	Value of relative air leaking at the piston position at top dead center (TDC), S_2	• Clearance between the ring and the 1-st compression ring by groove width, Y_1 • Clearance in gap of the 1-st compression ring. Y_2 • Clearance between the ring and the 2-nd compression ring by groove width, Y_3 • Clearance in gap of the 2-nd compression ring, Y_4 • Clearance between the piston and the engine block, Y_5 • Valve plug-to-guide bush clearance, Y_6 • Valve plug-to-guide bush clearance, Y_7
3	Flow rate of oil samp gas, S_3	• Clearance between the ring and the 1-st compression ring by groove width, Y_1 • Clearance in gap of the 1-st compression ring, Y_2 • Clearance between the ring and the 2-nd compression ring by groove width, Y_3 • Clearance in gap of the 2-nd compression ring, Y_4 • Clearance between the piston and the engine block, Y_5
4	Pressure in the main oil distributing passage, S_4	• Crank bearing-to-bushing clearance, Y_8 • Crankshaft neck-to-bushing clearance, Y_9 • Bush bearing of connection shaft-to-shaft neck clearance, Y_{10} • Bearing of distributive shaft-to-shaft neck clearance, Y_{11}

failures in the engine, causes of abnormal operations etc. Such information can be obtained during engine diagnostics by measuring the parameters that characterize the engine condition and comparing them with the standard values.

The variety and significant number of diagnostic parameters that describe the internal combustion engine condition necessitates the selection of the most informative ones that are characterized by the sensitivity of changes in their values depending on the changes of structural parameters and by unambiguous diagnostics. The set of diagnostic parameters of engine technicalcondition evaluation was substantiated based on the analysis of structural components failures and malfunctions statistics, trends of changes in the technical condition of mechanisms and units, developed structural diagrams of primary engine systems that define and limit its lifetime (cylinder-piston group, crank mechanism, valve timing gear). Equally important condition for the choice of diagnostic parameters is the ability to evaluate the engine remaining lifetime using their current values.

Therefore, such parameters as the analysis of the qualitative and quantitative composition of wear particles in oil, fuel burn rate, uncharging in combustion chamber, content of harmful substances in the exhaust fumes and others are not informative or require significant time of diagnosis. Table 3 shows the diagnostic parameters that meet the requirements, as well as a list of structural parameters that they evaluate.

Table 4 shows the standard nominal and limit values of diagnostic parameters specified by the manufacturer for the ZMZ-4061.10, 4063.10, 40637.10 engines.

Operational tests of engines dependability involved the effect of the clearance values given in Table 3 on the diagnostic parameters that evaluate them. For this purpose, before the second maintenance, the technical condition of the mechanical system of the internal combustion engine

Table 4. Standard values of diagnostic parameters assessing the engine condition

no.	Diagnostic parameter	Nominal value	Limit value
1	Pressure at the end of a compression stroke, kp/cm^2	12	9.6
2	Value of relative air leaking at the piston position at top dead center (TDC), kp/cm^2 for at least 5 sec	decrease from 1.5 to 1	decrease from 1.5 to 0.75
3	Flow rate of oil samp gas at 4000 min^{-1}, for at least l/min	22	62
4	Value of pressure in the main oil distributing passage, kp/cm^2: at 2500 min^{-1} at 700-800 min^{-1}	 5.0 –	 3.0 1.1

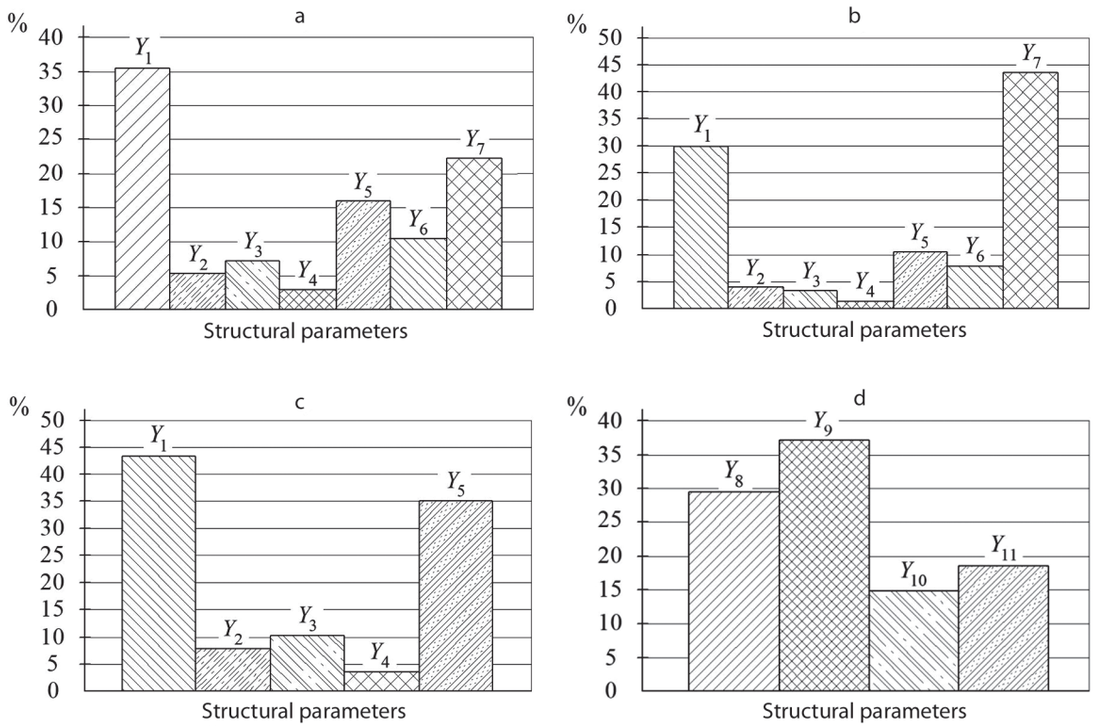


Figure 2. The degree of the effect of structural parameters on the diagnostic parameters: a) pressure at the end of a compression stroke; b)relative air leaking; c) flow rate of oil samp gas; d) pressure in the main oil distributing passage

was diagnosed. In case when values of diagnostic parameters exceeded the maximum permissible ones, the engines were sent to the repair department where they were partially or, if necessary, completely disassembled and the corresponding clearances measured. The collected measurement data was organized in a database and based on the results of its processing regressive models were built that characterized the effect of the structural parameters Y of the engine's systems on the diagnostic parameters S chosen for their evaluation:

$S_1 = 0.355\ Y_1 + 0.054\ Y_2 + 0.073\ Y_3 + 0.031\ Y_4 + 0.16\ Y_5 + 0.105\ Y_6 + 0.223\ Y_7$;

$S_2 = 0.298\ Y_1 + 0.037\ Y_2 + 0.033\ Y_3 + 0.015\ Y_4 + 0.104\ Y_5 + 0.077\ Y_6 + 0.436\ Y_7$;

$S_3 = 0.434\ Y_1 + 0.078\ Y_2 + 0.103\ Y_3 + 0.035\ Y_4 + 0.350\ Y_5$;

$S_4 = 0.294\ Y_8 + 0.372\ Y_9 + 0.148\ Y_{10} + 0.186\ Y_{11}$.

The findings showed that the degree of the effect of the same structural parameters on the diagnostic parameters chosen for the evaluation of the technical condition of internal combustion engines has different values. For example, the clearance between the ring and the 1-st compression ring by groove width (Y_1) has a dominant effect on diagnostic parameters S_1, pressure at the end of a compression stroke (35.5%) and S_3, flow rate of oil samp gas (43.4%). The structural parameter Y_7, valve-to-valve seat (43.6%) has the greatest impact on the relative pressed air leaking S_2. The crankshaft neck-to-bushing clearance Y_9 (37.2%) and crank bearing-to-bushing clearance Y_8 (29.4%) has an effect on diagnostic parameter S_4, pressure in the main oil distributing passage. Figure 2 shows the degree of the effect of clearance on the diagnostic parameters (in percentage points) in the form of diagrams.

The obtained dependences between the diagnostic and structural parameters allow identifying the most probable failures of the engine's mechanical systems and making the required list and algorithm of technical measures to restore their operability. For example, if the diagnostic parameter S_2 (relative air leaking at the piston position at top dead center) is out of tolerances, it most likely indicates increased wear of deflation valves, pistons and compression rings and in to a lesser degree changes in other engine components.

In case of deviation from standard values of diagnostic parameter S_1 (pressure at the end of a compression stroke) the most probable malfunctions are wear in the piston-to-compression ring, valve-to-valve seat and piston-to-engine block systems. Out of tolerances diagnostic parameter S_4 (pressure in the main oil distributing passage) indicates wear in the crankshaft neck-to-bearing bushing, crank bearing-to-bottom-end bearing systems, as well as wear of the necks of connection and distributive shafts.

The findings regarding engine operational dependability allow optimizing the system of their maintenance and repair, developing an algorithm for identification and elimination of occurring malfunctions. For example, up to the operating time of 90 thous. km there is no need to verify the condition of an engine's mechanical systems, since the probability of their fail-safe operation is at a sufficiently high level. In the operation time interval between 90 and 122 thous. km, it is recommended to perform deep diagnostics of gas distribution mechanism couplings, as in this interval there

is a significant increase of failure probabilities of its components (deflation valve, hydraulic tappet, cylinder head gasket). Starting from the operation time of 122 thous. km the technical condition of all engine structural parameters must be diagnosed.

The application of the findings in the processes of maintenance and repair of vehicles allows increasing the operational dependability of engines and reducing the costs associated with insuring their operability.

References

[1]. Bazhenov YuV. Osnovy teorii nadezhnosti mashin: ucheb. posobie dlya vuzov [Foundations of the machines theory of dependability: a study guide for higher educational establishments]. Moscow: Nauka; 1971 [in Russian].

[2]. Bazhenov YuV, Bazhenov MYu. Prognozirovanie ostatochnogo resursa konstruktivnykh elementov avtomobiley v usloviyakh ekspluatatsii [Predicting the residual life of automobile structural components in operation]. Fundamentalnie issledovania 2014;8:18-23 [in Russian].

[3]. Bazhenov YuV, Kalenov VP. Obespechenie rabotosposobnogo sostoyaniya elektronnykh sistem upravleniya dvigatelem v ekspluatatsii [Ensuring the operability of the electronic engine control systems in operation]. Avtomobilnaya promyshlennost 2015;12:23-27 [in Russian].

[4]. Denisov AS, Kulakov AT. Obespechenie nadezhnosti avtotraktornykh dvigateley [Ensuring the dependability of tractor engines]. Saratov: SGTU Publishing; 2007 [in Russian].

[5]. Denisov IV, Smirnov AA. Research of the operational dependability of the Lada Kalina vehicle systems affecting traffic safety. Dependability 2017;4:31-35.

[6]. GOST 27578–87. Technical diagnostics. Diagnosis of products. General requirements. Moscow: Izdatelstvo standartov; 1988 [in Russian].

[7]. Metodicheskie oukazania. Nadezhnost' v tekhnike. Metody otsenki pokazateley nadezhnosti po eksperimental'nym dannym. RD 50-690-89 [Guidelines. Technology dependability. Methods of dependability indicators evaluation based on experimental data. RD 50-690-89]. Moscow: Izdatelstvo standartov; 1990 [in Russian].

About the authors

Yuri V. Bazhenov, Candidate of Engineering, Professor of Automotive Transportation, AG and NG Stoletov Vladimir State University, Vladimir, Russia

Mikhail V. Bazhenov, Candidate of Engineering, Associate Professor of Automotive Transportation, AG and NG Stoletov Vladimir State University, Vladimir, Russia, e-mail: bagenovyv@mail.ru

Received on 30.03.2018

Original article
Dependability, vol. 18 no.4, 2018 DOI: 10.21683/1729-2646-2018-18-4-28-35

What should mean dependability calculation of unique highly vital systems with regards to single-use mechanisms of spacecraft

Yuri P. Pokhabov, *Joint Stock Company NPO PM – Maloe konstruktorskoye buro, Zheleznogorsk, Krasnoyarsk Krai, Russia*

Yuri P. Pokhabov

Aim. *Calculations are an integral part of the development of any complex technical object. Normally, they are subdivided into the calculations to confirm product operability (kinematic, electrical, thermal, strength, hydraulic and pneumatic systems analysis, etc.) and calculations to confirm its dependability (calculation of reliability, longevity, maintainability, storability and other indicators). As it is understood and provided in statutory documents, dependability calculation involves procedures of identification of an object's dependability indicators using methods based on their calculation using reference information on the object's components dependability, on the dependability of analog objects, on the properties of the materials and other information available at the time of calculation. However, in the case of development of unique highly vital systems, obtaining statistical data for dependability calculation is impossible due to two conflicting conditions, i.e. the limited number of produced objects and the requirement of high accuracy of the input information. Nevertheless, in the author's opinion dependability calculations must be performed. The only question is how to calculate the dependability and what such calculation should mean.* **Methods.** *In the classic dependability theory, the conventional understanding of probability of no-failure is the frequency of failures in time, yet for unique highly vital systems the failure rate must tend to zero over the entire period of operation (preferably, there should be no failures at all). For this reason the concept of "failure" in the context of unique highly vital systems should probably be interpreted not as an event, i.e. any fact, which as a result of experience can occur or not occur, but as possible risk, i.e. an undesirable situation or circumstance that is characterized by the probability of occurrence and potentially negative consequences. Then, an event in the form of a real or potential failure in operation can be associated with a risk in the form of probability of failure with negative consequences, which in terms of the consequences is equally unacceptable with regard to unique highly vital systems. In this case dependability calculation can be reasonably substituted with risk assessment, a process that encompasses risk identification, risk analysis and comparative risk assessment. Thus, risk assessment enables the achievement of the target dependability directly by substantiating the stability of manifestation of a specific product's properties and not indirectly through undependability caused by failures of analog products.* **Results.** *The paper shows the procedure of risk assessment for unique highly vital systems. Using the example of a mechanical system with actuated parts represented by a spacecraft single-section pivoted rod the risk assessment procedures are shown. The feasibility of risk assessment with the use of design engineering analysis of dependability is demonstrated.* **Conclusions.** *It is shown that the absence of statistical data on the dependability of analogs of unique highly vital systems does not prevent dependability calculation in the form of risk assessment. Moreover, the results of such calculations can be a source and guidelines for adopting design and process engineering solutions in the development of products with target dependability indicators. However, legalizing the method of such calculations requires the modifications of the technical rules and regulations to allow for dependability calculation by other means than with the use of statistical data on the failures of analogs.*

Keywords: *unique highly vital system, calculation, dependability calculation, risk assessment, design engineering dependability analysis.*

For citation: *Pokhabov YuP. What should mean dependability calculation of unique highly vital systems with regards to single-use mechanisms of spacecraft. Dependability 2018;18(4): 28-35. DOI: 10.21683/1729-2646-2018-18-4-28-35*

Introduction

The development of any complex technical products is impossible without calculations, i.e. *establishment and calculation of required data* [1]. Calculations in the form of ***documents*** *that contain calculations of parameters and values, e.g. dimension chain calculation, strength calculation, etc.*, are part of the list of design documentation per GOST 2.102. The codes and forms of calculations for engineering products are defined in the OST 92-0290 industry standard. For instance, according to GOST 2.119 the calculations are in general subdivided into the calculations to confirm product operability (kinematic, electrical, thermal, strength, hydraulic and pneumatic systems analysis, etc.) and calculations to confirm its dependability (calculation of reliability, longevity, maintainability, storability and other indicators). In technical rules and regulations (GOST 27.301 and GOST 27.410) dependability calculations are understood as only *procedures of identification of an object's dependability indicators using methods based on their calculation using reference information on the object's components dependability, on the dependability of analog objects, on the properties of the materials and other information available at the time of calculation.* Importantly, the availability of dependability calculations based on reference data on the dependability of analogs involves legal and financial implications in the context of insurance of the risks of loss of objects [2]. However, in the case of development of unique highly vital systems, obtaining statistical data for dependability calculation is impossible due to two conflicting conditions, i.e. the limited number of produced objects and the requirement of high accuracy of the input information. Despite the opinion that there is no need for dependability calculations for failsafe systems and they should be substituted with ensuring compliance with qualitative criteria of dependability [3], in the author's opinion dependability calculations as part of UHVC development are not optional. The only question is how to calculate dependability and what such calculation should mean.

The relevance of UHVC dependability calculation can be observed using the example of operation of single-use mechanisms of spacecraft. The efficiency of spacecraft operation in orbit wholly depends on the successful deployment of the solar panels and space antennas (reflectors), whose cost accounts for a negligible part of the total cost of the spacecraft and its placing into orbit. Experimental confirmation of dependable deployment is impossible due to high reliability requirements (0.9995 and higher) and unique environmental conditions of the deployment in orbit that cannot be accurately replicated as part of ground-based experimental activities. At the same time, practically any error in the design and manufacture of the deployment mechanisms may cause a failure that can entail the loss of the spacecraft. Therefore in this case dependability is largely defined by the calculations.

Approaches to the dependability calculation

Since failures of UHVC cause losses far greater than the cost of their creation [4], the dependability is characterized by the reliability and is defined by the indicator of probability of no-failure (PNF), i.e., *the probability that within the specified operation time no failure of the object occurs* [5, 6]. In the classic dependability theory PNF is normally understood as the frequency of failures in time, yet for the UHVC the failure rate must in theory tend to zero over the entire period of operation (preferably, there should be no failures at all). For this reason the concept of "failure" in the context of UHVC should probably be interpreted not as an **event**, i.e. *any fact, which as a result of experience can occur or not occur* [7], but as possible **risk**, i.e. *an undesirable situation or circumstance that is characterized by the probability of occurrence and potentially negative consequences* [8]. For single-use mechanisms of spacecraft we should talk of the risk as the *effect of uncertainty on the goals*, where uncertainty is understood as *"the state of complete or partial absence of information required for the understanding of an event, its consequences and their probabilities"* [9]. Then, an **event** in the form of a real or potential failure in operation can be regarded as a **risk** (probability of failure with negative consequences), which in terms of the consequences for UHVC is equally unacceptable. In this case **dependability calculation** can with no damage to the meaning be substituted with **risk assessment**, a *process that encompasses risk identification, risk analysis and comparative risk assessment* [10]. Importantly, risks of failure have no aspect of frequency, yet the risk assessment allows predicting the development scenarios of undesired situations that may cause failures and using such estimates in the adoption of engineering solutions as part of the UHVC development process. Thus, risk assessment enables the achievement of the target dependability directly by substantiating the stability of manifestation of a specific product's properties [11] and not indirectly through undependability caused by failures of analogs [12].

The departure from the understanding of an "event" as a fact of *disturbance of an object's operability* [5, 6] in the context of dependability calculation gives sensitivity to the concept of "dependability" in terms of its terminological definition. In the author's opinion, the shift of the standard definition of the term "dependability" to the functional interpretation diverts from an understanding of dependability other than that adopted in the current mathematics of the dependability theory. The use of the concept of "function" in the terminological definition of dependability as the requirements *established in the regulatory, design, project, contract and* ***other documentation*** *for an object* [6] causes the abstraction of the physical processes occurring within products and consequently does not encourage risk analysis. For example, in the organizational and engineering documentation, the

deployment of folding spacecraft in orbit is considered as a function that enables the spacecraft's preparation to operation within the specified service life, but at the physical level it is achieved through planned and consistent operation of a set of design components that enable the performance of such function. The functional definition of dependability actually makes "invisible" the operation of structural components that ultimately ensure dependable performance of the function of deployment of spacecraft's folding structures.

In the author's opinion, the definition of dependability as the *property of a system to maintain in time and within the set limits the values of all parameters* ***and/or indicators*** *that characterize the system's ability to perform the required functions in specified modes and conditions of operation, maintenance, storage and transportation* [13] provides a uniform understanding (self-consistency) of the parametric and functional definition of dependability [14] and enables dependability estimation both in terms of the classic dependability theory and in terms of analysis of the risks of failure. This becomes doable due to the fact that it is now possibility to consider dependability as a physical value with intrinsic simple and/or essential properties that can be expressed in parametric or non-parametric models through parameters and/or indicators [11].

The method of analysis of risks related to UHVC failures is based on the principles of physicality (causal connections) and physical necessity (consistency with the laws of nature) of the causes of failures. The task of the risk analysis while using the above principles becomes the analysis and synthesis of the simple properties that make the (essential) property of dependability, which becomes possible in the context of A.I. Uiomov's paradigm of the triunity of things, their properties and relations [15] and extended interpretation of the concept of "relation" as the mutual spatial arrangement, interrelation and interaction of things [11]. The connection between the parametric and non-parametric nature of properties' manifestation becomes evident if the term "operation" given in the now obsolete GOST 22487 standard is used, i.e. *"execution in the object (system) of a process (processes) according to the specified algorithm and (or) manifestation of specified properties by the object"*. In this case functions prescribed by the organizational and engineering documentation [6] at the physical level can be represented as the manifestation by an object of the specified properties in accordance with the specified algorithm of the performed process. This circumstance is extremely important in the context of technical systems, where during operation a number of properties can manifest themselves simultaneously of sequentially causing performance or non-performance of the functions specified in the organizational and engineering documentation.

This approach extends the capabilities of the classic dependability theory that is applied in strength calculations of dependability enabling additional evaluation of products' operation based on the mechanical, kinematic, energy, electrical and other parameters [16]. As at some hierarchical level physical properties are independent (e.g. the properties of strength and electrical conductivity), when examining any of the properties identified by the risk analysis it becomes possible to use either the deterministic or stochastic approach in the quantitative estimation of a specific dependability property under consideration.

Unlike classic dependability calculation, risk assessment enables the elimination of ambiguity in the product development process, i.e. taking into consideration the fact that the designer's idea must be reflected in the design documentation in a way that ensures that this idea is clear to the persons not involved in the design process and not familiar with the original ideas without additional explanations and comments and most importantly without the loss of meaning. Typically, ambiguity stems from the perception of the term "operable state" that is defined as the *state of an object in which it is able to perform the required functions* [6]. Taking into account the explanation of the understanding of the concept of function in the term "dependability" given in the national standard, it is not possible to qualify the operable state as sufficient for the performance of a product's intended function. The situation is somewhat clarified by the explanation of the term "operable state", according to which it can be defined as a *state of an object in which the values of all parameters that characterize the ability to perform the specified functions comply with the requirements of the documentation for such object* [6]. This certainly is a more specific definition of the operable state for a complex technical object, but it also has serious inaccuracies. First, for UHVC the requirement in the documentation must be necessary and sufficient, but the national standard does not clarify how to achieve that, which undoubtedly increases the role of the human factor in the development process (some people believe that the requirements are sufficient for achieving the object's operable state, some people don't). Second, the primary document for the products' manufacture is the design documentation and not *another documentation*, as the same standard puts it. In this sense the abandonment of the previous definition of the term "operable state" [5] that clearly specified design documentation in no way contributes to the reduction of the role of the human factor (due to less precise definitions).

Example and sequence of risk assessment

Let us examine an example of risk assessment in its standard form. In accordance with the definition of the term "risk assessment" [10], at the first stage the *risk identification* is performed, which consists in the identification of the source of risk and possible causes of failure. At this stage the product functionality is identified at the physi-

cal level in accordance with GOST 28806 in the form of *availability and specific properties of a set of functions capable of satisfying the specified or assumed needs*. The aim of this procedure is to provide the formal description of failures as hypothetical situations that prevent the performance of the functions under consideration. It is assumed that each potential failure is due to causes that directly engender them, that appear, exist and develop within the conditions of the environment as a set of external factors and operating modes in view of the worst possible combinations. Obviously, each type of failure can have several causes at once. The identified possible causes of failures as a whole are the foundation of a check list of risk identification. It must be understood that the risk identification procedures define the completeness of the identified object functionality and must be performed by qualified experts, as the results of such procedures fill-up the check list and ultimately serve as the criteria for the establishment of the obligatory and sufficient requirements in the design documentation.

At the next stage of risk assessment analysis is performed that generates the information background for the comparative risk assessment and adoption of decision regarding their sources. The procedures of risk analysis follow a specific algorithm in strict compliance with the general logic of actions according to the check list (in this case the identified causes of failure are the starting point for any subsequent actions related to risk analysis and assessment):

• properties of the critical components are identified, whose presence makes each cause of failures impossible,

• each property of critical components is defined quantitatively based on parameters (indicators),

• for each parameter (indicator) a range of allowed values is defined based on the requirements of the design specifications (the customer's idea of the product) and product build (the developer's idea of the product design),

• the value of each parameter within the allowed range is substantiated by calculations and experiments in terms of operability and dependability,

• dependability is evaluated by method of dependability structure diagram in order to confirm the fact that the selected values of the parameters (indicators) comply with the specification requirements,

• operability conditions are verified for parameter values compliance with the requirements of the norms, specifications and design documentation (for each parameter there must be a corresponding requirement for manufacture and/or operation, whose performance can be verified by means of maintenance inspection),

• risks are identified that are associated with failures as the result of absence of requirements in the detailed design and process engineering documentation "as is".

• probability is analyzed of failures associated with the underestimation of design and/or process engineering errors made during the development of the detailed documentation for adoption of the final decision on the compliance of the design and detailed documentation with the specified dependability requirements.

At the final stage of risk assessment the value of identified probability of failure is compared with the specified reliability requirements and, if necessary, actions are taken to reconsider the engineering solutions and/or establishment of additional requirements in the detailed documentation.

An example of risk assessment

As a specific example of risk assessment let us examine a mechanical system with actuated parts represented by a spacecraft single-section pivoted rod that for some time is fixed on the resting surface with a locking device, then the mechanical constraints in the lock are removed, the rod, by the action of actuators, is deployed to the specified angle, locks in the end position and starts operating as a panel with specified performance parameters [4]. The reliability of rod operation is ensured by sequential performance by its structural components of their assigned functions that consist in the manifestation of the strength of the rod under load in the locked position, prevention of spontaneous removal of mechanical constraints in the lock, transmission of electrical signal to the electric fuses of pyro cartridges upon command, pyro cartridge firing, removal of mechanical constraints in the lock, separation of the rod from the resting surface, rod rotation through the specified angle, locking and specified operation of the rod in the service position. The structural components of the rotating rod during deployment must sequentially perform all of the above functions in the assigned conditions and modes of operation. Failure to deploy the rod may be due to the failure of any of the functions or a combination of causes that may be defined not so much by the conditions and modes of operation as a combination of adverse factors.

As an example, let us examine the function of rod rotation through the specified angle with the deployment actuator. Failure of the above functions may be caused by the following conditions: non-activation or breakdown of the actuator (failure to activate), absence of required reserve of drive moment (deceleration), disappearance of radial clearance in the joint (joint locking), disappearance of axial clearance in the articulated joint (wedging), sudden appearance of obstacles in the rod's path (catching).

Obviously, each of the causes of failures can be countered by solutions and/or actions of the rod developer that provide its design with such critical component properties that would enable unconditional fulfillment of the assigned functions. For instance, to prevent or attenuate the consequences of:

• failure to activate the actuator it is required to ensure the limit probability of its faultless operation by means of redundancy of critical components,

• deceleration of the rod, creating a sufficient reserve of drive moment relative to the moment of resistance forces in its path by selecting the correct power performance of the actuator,

• joint locking. Choosing such radial clearances in the bearing as to ensure rotation freedom subject to possible changes in the thickness of the layer of solid lubricant and thermal deformation,

• wedging in the articulated joint, making provisions for thermal decoupling in the direction of the bearing's axis of rotation,

• catching of the rod, eliminating all possible obstacles in the rod's path caused by the gravity-free environment, kinematics of the motion or design of adjacent structures.

The quantitative estimation of the conditions of operability per each identified property of critical components involves choosing a parameter (indicator) that fully characterizes the property in question and the corresponding allowed range of deviation [16]. The range of allowed deviation of the parameters (indicators) will be defined by the requirements of the design specifications (external parameters) or by the internal design parameters (selected materials, layout and force diagrams, manufacturing processes, etc.) [17].

Let us cite the parameters (indicators) and their allowed ranges that correspond to the unconditional fulfillment of the function of rod rotation through the specified angle in the form of conditions that prevent or attenuate the consequences of the causes of failures for the following risks under consideration:

1) failure to activate the actuator

$$\mathrm{P}_d \geq P_{lim}, \tag{1}$$

where P_d is the probability of activation (operation) of actuator; P_{lim} is the probability of fault-free operation of actuator in accordance with the distribution of the assigned requirement of rod dependability indicator per structural components

2) deceleration of the rod

$$M_d > M_c, \tag{2}$$

where M_d is the drive moment developed by the rod deployment actuator; M_c is moment of resistance forces in the rod's path

3) joint locking

$$\Delta_r = д - 2д_n - д_{pr} > 0, \tag{3}$$

where Δ_r is the radial clearance in the joint; д is the minimum clearance in the connection between the internal and external members of the joint not including the layer of lubricant; $д_n$ is the maximum thickness of solid lubricant subject to its possible changes in the course of operation; $д_{pr}$ is the limiting value of thermal deformations in the radial clearance in case of volume expansion (compression) of the internal (external) member of the joint

4) wedging in the articulated joint

$$\Delta_{sh} > \Delta l, \tag{4}$$

where Δ_{sh} is the axial clearance in the articulated joint; Δl is the thermal deformation, capable of causing thrust force within the articulated joint

5) catching of the rod

$$Q_{st} \to 0, \tag{5}$$

where Q_{st} is the probability of the rod being caught.

The fulfillment of each of the conditions (1) to (5) in the course of operation under the given conditions and modes can be expressed in the form the probabilities that the values of the parameters (indicators) do not exceed the allowed limits over the observation interval t and will equal

$$P_1(t) = P(P_d \geq P_{lim}), \tag{6}$$

$$P_2(t) = P(M_d > M_s), \tag{7}$$

$$P_3(t) = P(\Delta_r > 0), \tag{8}$$

$$P_4(t) = P(\Delta_{sh} > \Delta l), \tag{9}$$

$$P_5(t) = 1 - Q_{st}. \tag{10}$$

The probabilities (6) to (10) can be identified by stochastic or deterministic methods. In the first case, the probabilities of parameters being within the allowed range are calculated using the methods of the dependability theory, e.g. method of individual dependability [18] (which ultimately does not rule out possible failures, but can provide the idea of their possible frequency). In the second case the fact of the parameters being within the specified allowed range is substantiated (necessary measures are taken to prevent failures) based on the provision of design reserves (redundancy, safety factor, drive moment reserve, parametric redundancy, power and temperature decouplings, procedures to ensure guaranteed results, e.g. by using minimax criteria).

Under the deterministic approach, in order to achieve the probabilities $P_i(t) \approx 1$, where $i = 1, 2, \ldots, 5$, in expression (1) the actuator must be redundant, e.g. for an electromechanical actuator a redundant motor power supply must be provided, while for a mechanical actuator structural redundancy must be in place; in expression (2) it is required to ensure drive moment reserves not less than 200 % for the worst combination of operating conditions and zero kinetic energy of the rod [19]; in expression (3) minimax criteria must be provided that are based on the restriction of the ranges of realization of random parameters for the worst conditions of

their realization [20]; in expression (4), thermal decouplings must be in place [21]; in expression (5), procedures to ensure guaranteed results are to be provided, e.g. with the use of computer simulation [22].

In case of application of any method (stochastic or deterministic) of probabilities (6)-(10) calculation the dependable performance of the function of rod rotation through the specified angle is identified using formula

$$P(t) = \prod_{i=1}^{n} P_i(t), \quad (11)$$

where n is the number of indicators that ensure unconditional fulfillment of the function of rod rotation through a specific angle; $P_i(t)$ is the probability of the i-th parameter (i = 1, 2, …, 5) not exceeding the allowed limits; t is the observation interval.

The calculated value (11) provides the theoretical dependability indicator that may differ from the real one if the design and/or process engineering documentation does not contain some manufacturing requirements or they are specified incorrectly for non-ambiguous fulfillment of conditions (1)-(5). The absence, ambiguity or incorrect performance of requirements of the technical documentation can be caused by events associated with failure to conduct the required calculations and tests, omissions on the part of designers in the preparation of drawings, limited time of delivery of design documentation, lack of coordination between designers and process engineers, etc.

In order to reduce the risks associated with the failures caused by insufficient scope or ambiguity of the requirements, the design and process engineering documentation must be analyzed for compliance of the scope of the parameters (indicators) that describe the performance of certain functions, e.g. (1)-(5), with the respective requirements.

Non-relevance of the parameters and requirements of the design and/or process engineering documentation, risks of non-fulfillment or undue fulfillment of requirements in the process of manufacture are regarded as events C_i, where index i corresponds to the i-th component of the system under consideration. The probability of each such event may be defined by formula:

$$P(C_i) = б_i \cdot P_i(t), \quad (12)$$

where $б_i$ are adjusting factors that can be obtained by expert methods, e.g. using point-based estimation of failure severity:

$$б_i = 1 - Q_i,$$

where Q_i is the expected probability of failure of the i-th component in accordance with the scale of point-based estimation of failure severity per GOST 27.310.

In order to calculate the final probability of the performance of the function of rod rotation through the specified corner subject to the provisions of the design and process engineering documentation (12) the following formula is used

$$P(C) = \prod_{i=1}^{n} P_i(C_i). \quad (13)$$

The above procedures of evaluation of the probability of performance of the function of rod rotation through the specified corner can be used as part of the analysis of each of the mentioned rod functions during deployment, while the probability of their performance and the general probability of no-failure of the rod are evaluated using formulas (11) and/or (13). The applicability of the above formulas is defined by the required accuracy of dependability evaluation [16]. For the purpose of estimation of reliability below three nines formula (11) may prove to be quite applicable, while if the required reliability is three nines and above formula (13) must be used.

Risk assessment with the use of design engineering analysis of dependability

The method and risk analysis and assessment subject to design and technical solutions (1)-(13) was named design engineering analysis of dependability (DEAD), whose general description is given in [23, 24]. The methodology can be described as a sequential performance of a set of specific methods:

• The functional analysis method is intended for the identification of the primary functions that enable the performance of products' intended function and identification of possible failures as the result of violation of operational conditions.

• Method of worst case analysis for the identification of the causes for possible failures including the worse combinations of factors of a product's technical condition, modes and conditions for its operation.

• The method of failure management for the identification of the properties of products' critical components, whose implementation makes the causes of failures impossible.

• Method of product design parametrization for quantification of the properties of critical components and definition of the ranges of allowed values, e.g. (1)-(5).

• Method of parameters substantiation for the evaluation of the probability of the operating parameters being within the allowed range, e.g. (6)-(10).

• Method of dependability evaluation using the method of dependability structure diagram (11) for decision-making regarding the compliance of the chosen design parameters with the assigned dependability requirements.

• Method of definition of necessary and sufficient requirements by means of continuous analysis of the design and process engineering documentation for identification of the degree of compliance of the operating parameters with the specified requirements.

• Method of identification of risks of failure due to non-specified requirements in the design and/or process engineering documentation (12) for identification and evaluation of possible failures as the result of compliance with the requirements of detailed documentation "as is".

• Method of dependability evaluation subject to the risks associated with the underestimation of design and/or process engineering errors (if identified), e.g. with the use of point-based estimation of failure severity (13) for adoption of final decisions regarding the compliance of the design with the specified dependability requirements.

Depending on the required accuracy of dependability estimation the obtained values of probabilities (11) or (13) are compared with the specified dependability requirements P_{pr} to ensure the fulfillment of condition

$$\forall P = [P(t) \vee P(C)] > P_{pr}. \quad (14)$$

In case of non-fulfillment of condition (14) DEAD procedures must be reiterated and new calculations must be performed with refined initial data.

It should be noted that the above approach to dependability calculation (11)-(14) was developed specifically for folding structures of spacecraft and has not yet been applied to other technical objects. Nevertheless, if we compare this approach with the procedure of dependability calculation of mechanical parts of aircraft rotary structures based on conventional approaches of the dependability theory [25, 26], the former allows significantly extending the capabilities of taking uncertainty factors into account. For example, out of five causes of failures considered in this paper, known sources only examine one, i.e. "rod deceleration" (2), which is completely explainable as such sources did not regard the design and process engineering solutions as uncertainty factors.

Conclusion

It is shown that the absence of statistical data on the dependability of UHVC analogs does not prevent dependability calculation in the form of risk assessment. Moreover, the results of such calculations can be a source and guidelines for adoption of design and process engineering solutions in the development of products with target dependability indicators. However, legalizing the method of such calculations requires the modifications of the technical rules and regulations to allow for dependability calculation by other means than with the use of statistical data on the failures of analogs.

References

[1]. Ushakov D. Dictionary of the Russian language in 4 volumes. Volume 3. Moscow: Terra; 1996 [in Russian].

[2]. Terminy kosmicheskogo strakhovaniya [Terms of space insurance], <http://www.space-ins.ru/index.php/kategoria8/8-terms.html> [in Russin].

[3]. Polovko AM, Gurov SV. Osnovy teorii nadiozhnosti [Introduction into the dependability theory]. Saint Petersburg: BHV-Peterburg; 2006 [in Russian].

[4]. Pokhabov YuP, Ushakov IA. O bezavariynosti funktsionirovaniya unikal'nykh vysokootvetstvennykh sistem [On the fail-safety of unique highly vital systems]. Metodi menedzhmenta kachestva 2014;11:50-56 [in Russian].

[5]. GOST 27.002-89. Industrial product dependability. General principles. Terms and definitions. Moscow: Izdatelstvo Standartov 1990 [in Russian].

[6]. GOST 27.002-2015. Dependability in technics. Terms and definitions. Moscow: Standartinform 2016 [in Russian].

[7]. Wentzel ES. Probability theory. Moscow: Nauka; 1969 [in Russian].

[8]. GOST R EN 9100-2011. Quality management systems of organizations of aviation, space and defence industries. Requirements. Moscow: Standartinform; 2012 [in Russian].

[9]. GOST R ISO 31000-2010. Risk management. Principles and guidelines. Moscow: Standartinform; 2012 [in Russian].

[10]. GOST R 51897-2011. Risk management. Terms and definition. Moscow: Standartinform; 2012 [in Russian].

[11]. Pokhabov YuP. About the philosophical aspect of reliability exemplified by unique mission critical systems. Dependability 2015;3:22-27.

[12]. Van-Jelen V. Fizicheskaya teoriya nadyozhnosti [Physical theory of dependability]. Simferopol: Krym; 1998 [in Russian].

[13]. Pokhabov YuP. On the definition of the term "dependability". Dependability, 2017;17(1):4-10.

[14]. Netes VA, Tarasyev YuI, Shper VL. How we should define what "dependability" is. Dependability 2014;4:15-26.

[15]. Uiomov AI. Veshchi, svoystva i otnosheniya [Things, properties and relations]. Moscow: USSR AS Publishing; 1963 [in Russian].

[16]. Pokhabov YuP. Teoriya i praktika obespecheniya nadyozhnosti mekhanicheskikh ustroystv odnorazovogo srabatyvaniya [Theory and practice of dependability of single-use mechanical devices]. Krasnoyarsk: SFU Publishing; 2018 [in Russian].

[17]. Chebotariov VE. Proektirovanie kosmicheskikh apparatov sistem informatsionnogo obespecheniya: v 2-kh kn. Kn. 2. Vnutrennee proektirovanie kosmicheskogo apparata [Design of information support spacecraft in 2 volumes. Volume 2. Internal design of spacecraft]. Krasnoyarsk: SibSAU Publishing; 2005 [in Russian].

[18]. Timashev SA, Pokhabov YuP. Problemy kompleksnogo analiza i otsenki individualnoy konstruktsionnoy nadyozhnosti kosmicheskikh apparatov (na primere povorotnykh konstruktsiy) [Problems of comprehensive analysis and assessment of individual design dependability of spacecraft (with the example of rotating structures)]. Ekaterinburg: AMB; 2018 [in Russian].

[19]. NASA Standard. Design and Development Requirements for Mechanisms; June 13, 2006. NASA-STD-5017.

[20]. Chebotariov VE, Kosenko VE. Osnovy proektirovania kosmicheskikh apparatov sistem informatsionnogo obespecheniya [Introduction to the design of information support spacecraft]. Krasnoyarsk: SibSAU Publishing; 2011 [in Russian].

[21]. Pokhabov YuP, Grinevich VV. Patent 2230945 Russian Federation. MPK F16B 1/00. Method of product fastening. No. 2002113143/11. Appl. 18.05.2002. Publ. 20.06.2004. Bull. No. 17 [in Russian].

[22]. Emelianov AA. Put ot analogovykh modeley k simulyatoru na tsifrovom kompyutere [From analogue models to digital computer simulator]. Prikladnaya informatika 2007;5:41-53 [in Russian].

[23]. Pokhabov Yu.P. Approach to ensuring of dependability of unique safety critical systems exemplified by large flexible structures. Dependability 2016;1:31-36.

[24]. Pokhabov YuP. Ensuring dependability of unique highly vital systems. Dependability 2017;17(3):17-23.

[25]. Kuznetsov AA. Nadyozhnost konstruktsii ballisticheskikh raket [Structural dependability of ballistic missiles]. Moscow: Mashinostroenie; 1978 [in Russian].

[26]. Kuznetsov AA, Zolotov AA, Komyagin VA et al. Nadyozhnost mekhanicheskikh chastey konstruktsii letatelnykh apparatov [Dependability of mechanical parts of aircraft design]. Moscow: Mashinostroenie; 1979 [in Russian].

About the author

Yuri P. Pokhabov, Candidate of Engineering, NPO PM – Maloe konstruktorskoye buro (OAO NPO PM MKB), Head of Center of Research and Development, Russia, Krasnoyarsk Krai, Zheleznogorsk, e-mail: pokhabov_yury@mail.ru

Received on: 16.04.2018

Original article
Dependability, vol. 18 no.4, 2018 DOI: 10.21683/1729-2646-2018-18-4-36-44

Method of assessing the protection of computer-based control systems under information technology interference

Sergey M. Klimov, *4th Central Research and Design Institute of the Ministry of Defence of Russia, Korolyov, Russia*
Yuri V. Sosnovsky, *Physics and Technology Institute, V.I. Vernadsky Crimean Federal University, Simferopol, Russia*

Sergey M. Klimov

Yuri V. Sosnovsky

Abstract. *The aim of this paper is to develop models that would enable a standardized representation of the structure, functions of computer-based control systems (CBCS) and quantification of the risk (fault tolerance) of automated control systems and their primary components, i.e. CBCS, under information technology interference (ITI). The paper shows the relevance and importance of CBCS models and estimation of the risk of operation of automated process control systems (APCS) under various ITI (computer attacks). Intruder ITI under consideration includes hardware, firmware and software-based interference able of blocking communication channels, disrupting information availability and integrity, as well as targeted and lasting information technology interference with an automated system, namely with the use of malware. The structural and functional model of a computer-based control system as the primary component of a higher-level system (APCS) developed in this paper is composed of a set of diagrams and descriptions of functions. The structural and functional model includes the following: channel structure of the control system's main cycle (reading, processing of data, recording of output values, as well as communication subsystem operations), structural and functional diagram of CBCS of various types depending on the availability and utilization of a communication channel within the structure of the control cycles, standard vulnerability certificate. The diagrams detail the standard functions, operating procedures and information interaction of CBCS modules with the environment via communication channels. The ITI-specific risk model of APCS and CBCS as its part is described by indicators that characterize the conditional harm and condition of the control system, in which it is able to recover its operability, or whether external intervention is required that would affect not only the control system itself, but the controlled process as well. The following indicators were examined: characteristic points and parameters of risk function based on the Weibull-Gnedenko distribution, statistical estimation of CBCS protection, risk function, dynamic estimation of the risk of successful implementation of ITI against CBCS. It is assumed that the values of the parameters required for the calculation of the risk parameters and CBCS protection were obtained:*
- empirically based on structural and parametric analysis of the design features, functional dynamics and vulnerabilities of CBCS
- as part of testbed simulation of CBCS as computer network users under ITI
- experimentally based on the frequency of successful ITI threats,
and the protection indicators are also extrapolated to the whole CBCS lifecycle by means of a dynamic risk function-based correction using the Weibull-Gnedenko distribution.
In the conclusion it is noted that the developed method of assessment of CBCS protection under ITI allows evaluating the risks of successful implementation by an intruder of malicious actions against CBCS and APCS in general, which predetermines the requirement for timely elimination of CBCS vulnerabilities and adoption of additional organizational and technical measures aimed at improving information security of automated control systems.

Keywords: *information technology interference, computer-based control systems, information protection facilities, fault tolerance.*

For citation: *Klimov SM, Sosnovsky YuV. Method of assessing the protection of computer-based control systems under information technology interference. Dependability 2018;4: 36-44. DOI: 10.21683/1729-2646-2018-18-4-36-44*

Introduction

The Doctrine of Information Security of the Russian Federation approved by order of the President of the Russian Federation in 2016 defines the current threats of information technology interference against the nation's critical information infrastructure.

Today, the necessities of the nation's developing digital economy define the active introduction of information and communication technologies as part of automated process control systems (APCS). Worldwide, APCS are classified as SCADA systems for control of power, transportation and industrial systems. In practice, the deployment of information and communication technologies causes the emergence of additional vulnerabilities in software, which increases the probability of realization of information technology interference (ITI) against them by an intruder.

Today's ITI malware [1], e.g. Stuxnet, Flame, miniFlame, Duqu, Gauss, Reign, Wiper, Shamoon, Careto exploit the vulnerabilities of the APCS software code for hidden deployment, self-propagation and intentional disruption of a system's operation. The development of ITI tools and their functional capabilities is significantly ahead of the corresponding tools of detection and prevention of computer attacks (CADPS), especially in the form of malicious software.

The key element of APCS CADPS is the sensor (firmware or software display module) that detects the fact of a computer incident, i.e. successful implementation of an ITI by an intruder.

The basic element of APCS are computer-based control systems (CBCS) [2], whose software performs the functions of collection, processing and transmission of information for the purpose of real-time control of critical facilities. While earlier programmable industrial microprocessors were controlled by means of sets of special commands, today they operate under the control of general-purpose operating systems (OS) (e.g. Windows or Linux) and are available as users of computer networks with TCP/IP data protocols or Modbus data protocol typical for SCADA systems.

The assessment and protection of information in CBCS affected by an intruder's ITI requires a system of methods and tools for detection, identification of malicious actions and elimination of their consequences [9-14].

Thus, the task of developing a method of evaluation of CBCS protection against ITI for the purpose of a priori quantitative estimation of the risks of violation of critical facilities' CBCS operation is of relevance and practical interest.

Problem definition

The research is based on the following premises:

- estimation of the risks of an intruder's successful ITI must be carried out using a testbench that allows creating the required test conditions for the operation of functional equivalents of CBCS, elements of CADPS and ITI simulation

- the identified groups of risks of CBCS operation disruptions can be assessed on site using a mobile test suite

- preliminary assessment of CBCS vulnerabilities and an intruder's ITI threats allows defining possible information protection facilities (IPF) and choose the most efficient ones.

The CBCS protection parameters are defined:

- empirically based on structural and parametric analysis of the design features, functional dynamics and vulnerabilities of CBCS

- experimentally based on the frequency of an intruder's successful ITI

the identified protection indicators are also extrapolated to the whole CBCS lifecycle by means of a dynamic risk function-based correction using the Weibull-Gnedenko distribution [3-4].

The development of a method of evaluation of CBCS protection under ITI is based on the model of CBCS operation under ITI that enables comprehensive analysis of mutually related processes of CBCS operation, ITI implementation and elimination of consequences.

The final model implies the CBCS is equipped with a communication subsystem. The communication subsystem enables such functions as interaction with industrial systems of a higher level, remote reading of sensor data and recording of their values into executive devices by means of network interfaces.

The CBCS communication subsystem that is directly included into the process control loop, is the primary vulnerability for the implementation by the intruder of the ITI threats against its data communication protocols. It is assumed that an intruder, while implementing an ITI, may exploit undocumented features of both the hardware and software facilities of a CBCS, and programmable routers of the data communication network at various APCS levels. Additionally, an intruder may be both on the outside and inside and be aware of the specificity and time limits of a process (triggering conditions of the automatic and automated executive devices) and is able to implement an unknown zero-day action.

The diagram of the model of CBCS operation under ITI in terms of an augmented Petri net (APN) [5] is shown in Figure 1.

Model of CBCS operation under ITI includes three loops:

1. Normal CBCS operation loop that enables simulation and structure and parameter analysis of the CBCS control loop (CL).

2. ITI simulation loop designed to simulate an intruder's actions associated with the acquisition of unauthorized access to CBCS, passive and active vulnerability scanning, selection and launching of ITI. For the input information regarding the modern threats of malware-based ITI this paper refers to [1].

3. ITI consequences elimination loop that enables the simulation of the processes of prevention, detection and

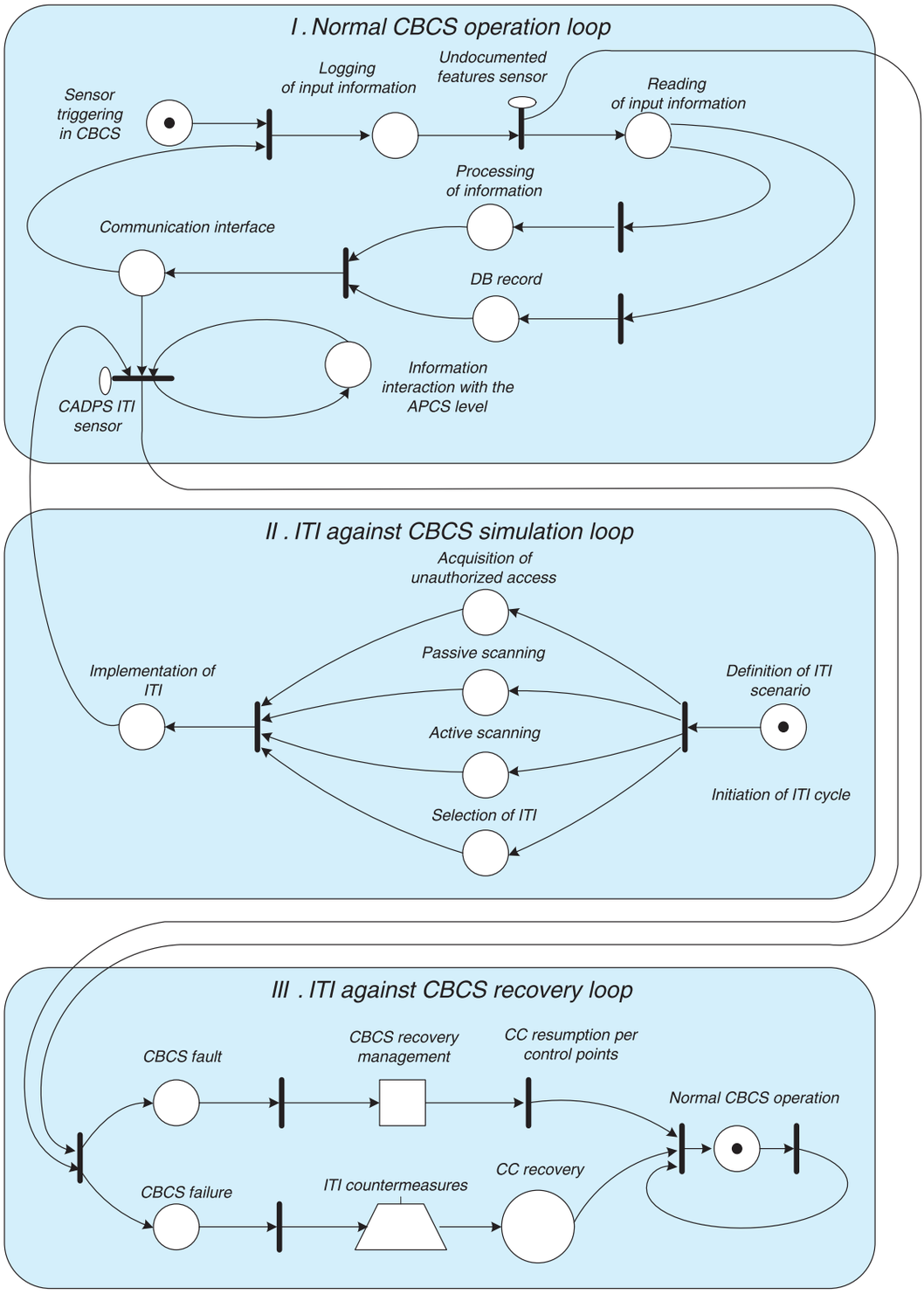


Figure 1. Diagram of the model of CBCS operation under ITI in terms of APN

elimination of ITI consequences based on the use of sensors (indication modules) of interference detection and identification.

In Figure 1, block III shows two alternative solutions for the elimination of ITI consequences for CBCS in case of their successful implementation by an intruder.

Branch I (upper) shows the situation in which a CBCS demonstrates the recovery after fault (relatively short CC, i.e. from several seconds to several minutes). A fault is understood as a short disruption of CC caused by ITI that yet does not entail CBCS failure. In this case, if a fault is identified, the system launches the CC recovery algorithm,

upon the completion of which the CBCS enters the state of normal operation.

Branch II (lower) shows the ITI implementation approach, under which the CBCS enters the state of failure that is characterized by long disruption of CBCS control processes (from 30 minutes to several hours).

The specificity of CBCS is such as the CC recovery after a long failure often requires the involvement of the operator and/or technical personnel and cannot be performed by means of a reset or deployment of an a priori operable CBCS.

Preventing faults and failures of CBCS under ITI requires prompt detection, localization of the interference and CC recovery based on the deployed redundant CADPS sensors (indication hardware and software facilities) [5-6].

Formalization of the model of CBCS operation under ITI and in terms of APN [5]:

$$S_{CBCS} = \{(P,V),T,D,M,Q,I_p,Y\}, \quad (1)$$

where $P = p_1, p_2, ..., p_i$ is a nonvacuous finite set of places that characterize the normal CBCS operation mode

$V = v_1, v_2, ..., v_i$ is the set of recovery places that reflect the procedures of recovery after an intruder's successful ITI (graphically presented as □)

$T = t_1, t_2, ..., t_i$ is a nonvacuous set of transitions. According to APN, each transition t_i can be associated with the triggering algorithm $a \vee g_i$ (if the algorithm is available the transition is marked with $a \vee g_i$)

D is a nonvacuous finite set of net arcs, while $D = (D_1 \cup D)$, $D_1 = (P \times T) \cup (V \times T)$ is a nonvacuous set of input arcs connecting places and transitions, $D_2 = (T \times P) \cup (T \times V)$ is a nonvacuous set of output arcs oriented from transitions to places

M is the set of Petri net markings

$F_p : (M_p : P \to N), F_v : (M_v : V \to N)$ are the functions of the initial marking of the places of normal operation and recovery, respectively, N={0,1,2,…} is a set of natural numbers (marked with a dot inside the place ⊙)

Q is the set of probabilities of transition firing that represents the probabilities of CBCS being in normal operation, moments of ITI implementation and CADPS sensors triggering, recovery processes

Z_{ACMi} is the set of places of ITI countermeasures ( )

$I_p = i_{p1}, i_{p2}, ..., i_{pm}$ is the set of priorities for arcs

$Y = y_1, y_2, ..., y_k$ is the set ITI temporal parameters.

The functions of description of APN structure in the form of set mapping are as follows:

$$F_{d1} : P \times T \cup V \times T \to N, \text{ or } F(p_i, v_j, t_n), \quad (2)$$

$$F_{d2} : T \times P \cup T \times V \to N, \text{ or } F(t_n, p_i, v_j),$$

where F_{d1} is the function of input places that associates the number of markings required for transition firing ("input") with the places and transitions

F_{d2} is the function of output places that associates the number of markings required for the modification of marking (correction of "output") with the places and transitions

N={0,1,2,…} is a set of natural numbers.

Given the above, the rule of transition firing has the following standard form:

$$\forall (p_i \in P \wedge v_j \in V) \to \exists (M(p_i) \geq F_{d1}(p_i, v_j, t_n)). \quad (3)$$

If transition t_n is triggered, out of each of its input places p_i and v_j the number of markings $m(p_i)$ and $m(v_j)$ is removed that is equal to the number of input arcs, while to the output places p_{i+1} and v_{j+1} the number of markings is added that is equal to the number of output arcs. The transition is triggered that corresponds with the highest probability of its firing (q_w) and is preceded with the arc with a higher priority (i_{pm}). The delay of transition triggering time is defined by the ITI parameters (y_k) in the net places connected to such transition with arcs. Accordingly, the APN marking change rule is as follows:

$$\forall (M_p \wedge M_v) : (p_i \in P \wedge v_j \in V) \to \exists (M_p' \wedge M_v') =$$
$$= F_p(Mp) + F_v(Mv) - F_{d1}(p_i, v_j, t_n) + F_{d2}(t_n, p_i, v_j). \quad (4)$$

Description of initial marking (M_i) in APN for presentation and analysis of causal relationships between processes in CBCS and CADPS under ITI. Condition of achieving APN:

$$\forall (p_i \in P \wedge V_r \in V \wedge Z_j \in Z_{ACM}) \to \exists M'(p_i, v_r, z_j) =$$
$$= M(p_i, v_r, z_j) - [F_p(M_p) + F_v(M_v) + F_z(M_z)] -$$
$$- [F_{ud1}(p_i, v_r, z_j, t_n) + F_{ud2}(t_n, p_i, v_r, z_j)]. \quad (5)$$

Definition of logical conditions of firing of APN transitions (T_i) under marking $M(p_i, v_r, z_j)$:

$$\forall (p_i \in P \wedge V_r \in V \wedge Z_j \in Z_{ACM}) \to$$
$$\to \exists [M(p_{i1}, v_{r1}, z_{j1}) \geq F_{ud1}(p_i, v_r, z_j, t_n)],$$
$$\Psi_t [(p_{i1}, v_{r1}, z_{j1}), ..., (p_{in}, v_{rn}, z_{jn})] = 1, \quad (6)$$

where Ψ_t is the function of marking distribution per APN input places.

Definition of relation for APN place, "subevents" P_i, V_r, Z_{ACMi} of ITI warning, detection, analysis, active countermeasures as well as CBCS recovery:

$$\forall (p_i \in P, e_i \in E_{ki}, n_i \in N_{ki}, b_i \in B_{ki}) \to$$
$$\to \exists \min (p_i, V_r, Z_j, E_{ki+1}, N_{ki+1}, B_{ki+1},) \to$$
$$\to Z^*_{SE} \to, \Psi_{tn} = \{(p_{i1}, V_{r1}, Z_{j1}), ..., (p_{in}, V_{rn}, Z_{jn})\}. \quad (7)$$

The function of the starting input distribution of marking Ψ_t over APN places takes the value (8) that defines the order of the starting allocation of markings to APN places

$$\Psi_{tn}\left(p_i, V_r, Z_j\right) = \left\{1,\ if\ m_{pi} \in M_p, m_{vr} \in M_v, m_{zj} \in M_z. \right. \quad (8)$$

Definition of CADPS sensor triggering conditions:

$$\forall t_i \in T, Q_{ki} \in Q, \exists Q_{i+1} \neq 0,$$
$$\phi_q\left(t_i, Q_{ki}\right) = \left\{1,\ if\ mp_i \in M_p, y_i \in Y, a \vee g_i = \right.$$
$$= 1, \left\{\frac{1}{R},\ if\ mp_i \in M_p, y_i \notin Y, a \vee g_i = 1, \right. \quad (9)$$

where $a \vee g_i$ is the sensor triggering algorithm. The sensor triggering conditions are as follows:

$\phi_q\left(t_i, Q_{ki}\right) = 1$, sensor triggered, attack detected

$\phi_q\left(t_i, Q_{ki}\right) = \frac{1}{R}$, false sensor triggering with the R-th transition triggering

$\phi_q\left(t_i, Q_{ki}\right) = 0$, sensor not triggered, unknown attack not detected.

Taking into account today's methods of information systems protection and risk management [1, 5, 7, 9, 13, 14], the method of assessment of CBCS protection under ITI is presented as the following sequence of steps:

1. Definition of the method of CBCS monitoring taking into account the particular operating principles (controlled processes and types of data communication protocol).

2. Analysis of vulnerabilities and generation of CBCS vulnerabilities certificate.

3. Development of the ITI threat model.

4. Development of the simulation model of the controlled CBCS CC taking into account the APCS communication interfaces.

5. Experimental research of CBCS under ITI based on testbed simulation.

6. Evaluation of CBCS protection indicators based on the simulation results.

7. Assessment of the risks of CBCS protection disruption under ITI.

Step1. CBCS monitoring is based on the classification of CBCS. CBCS classification is based on attributes that include the availability of wired and wireless communication channels, interfaces (unidirectional, bidirectional, multipoint), capability of remote firmware replacement and remote CBCS administration. Based on the above attributes, let us identify the basic CBCS types:

CBCS of the 1-st type, i.e. system with a local controller performing local reading of input signals, data processing and generation of output control signals by means of local output modules

CBCS of the 2-nd type, i.e. system that uses data communication interfaces as the information environment between remote input-output modules and processor units

CBCS of the 3-rd type, i.e. systems that use data communication protocols that implement two-way communication to transmit data to higher-level systems and receive data from them

CBCS of the 4-th type, i.e. systems that have the properties of the systems of the 3-rd type, but allow remote (using the common data communication environment) administration, including correction and change of the control program (firmware replacement).

The controlled processes include internal and external data exchange via CBCS interfaces, parameters of network traffic to APCS components. Data exchange monitoring in CBCS is divided by types of data communication protocols and is performed by CADPS sensors through signature analysis and functional analysis of abnormal CBCS behaviour, detection of distortions of protocol structure, signalling and synchronization parameters, data packet preambles and various service parameters of CBCS equipment.

Step 2. CBCS software vulnerabilities shall be identified based on the provisions of GOST R 56546-2015 [8] and formalized as a standard certificate of CBCS software vulnerabilities under ITI (Table 1). The BDU:2018-00543 database vulnerability certificate by the State Research and Design Experimental Institute of Technical Information Protection of the Federal Service for Technical and Export Control of Russia was taken as a model. The following characteristics must be introduced to compliment the standard vulnerability certificate:

1. Type of industrial data communication protocol (due to potential unique attacks against industrial protocols that take into consideration the specifics of their hardware and software design).

2. The vulnerability vector must include data on the controlled process (regular, important, critical), as information technology interference against CBCS, APCS causes not only the fault or failure of such system, but also the fault or failure of the controlled process.

3. The Vulnerability Hazard Level indicator depends on both the vulnerability threat level, and the type of the controlled process (regular, critical).

Step 3. The development of the ITI threat model is based on the analysis of the potential threats that depend on the type of the CBCS (per the classification), as well as the used protocols and control cycles (CC) under control.

Given that information security is examined at CBCS level, the following primary ways of ITI implementation can be identified:

- threats of information technology interference originating at a higher level of APCS (such threats may include incorrect settings for a specific process, their enforcement, other attempts of interfering with the CBCS operation through distortion of information and control parameters)
- threats of interference with the CBCS information input and output protocols (in the case of CBCS of the types 2 to 4) causing the blocking of the protocols or disruption of the integrity of transmitted data
- threats of interference with the software component of CBCS (in cases when remote firmware replacement is available).

An intruder's capability to implement ITI threats are directly associated with the characteristics of the employed equipment, protocols, controlled CBCS processes and cannot be considered out of this context.

Table 1. Standard vulnerabilities certificate (exemplified by BDU:2018-00543)

Vulnerability description	Vulnerability of the track_import_export.php scenario of the U.motion builder manufacturing and residential buildings management system is associated with non-adoption of measures for protection of SQL query
Vendor	Schneider Electric
SW name	U.motion Builder
SW version	up to 1.3.4
SW type	APCS software tool
OS and hardware platforms	TBD
Error type	Non-adoption of measures for protection of SQL query structure (SQL injection type of attack)
Error type identifier	CWE-89
Vulnerability class	Code vulnerability
Date of detection	02.03.2018
Underlying vulnerability vector	AV:N/AC:L/Au:N/C:C/I:C/A:C
Hazard level of vulnerability	Critical level (CVSS 2.0 base rate is 10) High level (CVSS 2.0 base rate is 8.8)
Possible vulnerability elimination measures	Application of recommendations given at https://www.schneider- electric.com/en/download/document/SE_UMOTION_BUILDER/
Vulnerability status	Confirmed by manufacturer
Presence of exploit	TBD
Method of exploitation	Injection
Method of elimination	Software update
Information on elimination	Vulnerability eliminated
Source reference	https://www.schneider-electric.com/en/download/document/SE_UMOTION_BUILDER/
Identifiers of other vulnerability description systems	CVE: CVE-2018-7765
Other information	

Step 4. Development of the simulation model of the controlled CC in CBCS (research object) subject to APCS communication interfaces consists in the development of the so-called "information environment" and basic CBCS functions. In the course of testbed experiments a CBCS model can be represented as a separate group of programmable devices, e.g. programmable logical switches, whose software enables CBCS CC modeling subject to changes in the intentional and non-intentional control and information actions. In the course of simulation the role of "information environment" is played by independent software blocks of the model, while in the course of full-scale simulation this role is played by real equipment involved in the CC control. As part of simulation, the real CBCS CC may remain inactive (due to the complexity of complete replication of APCS functions), while the simulation of its reactions to certain actions can be executed by adding independent software modules to the special software of the communication equipment.

Step 5. Experimental research of CBCS under ITI based on testbed simulation involves three primary processes:

- simulation of normal CBCS operation
- simulation of ITI against CBCS out of the database of interference models
- simulation of IPF (for the purpose of this paper, CADPS) configured for CC monitoring subject to the proposed CBCS classification.

The program code and ITI simulation scenarios are stored in the database and are developed subject to the employed protocols, types and potential vulnerabilities of CBCS (per the classification). Some ITI may include standard computer attacks, e.g. ARP spoofing, DDoS, while others are specific to the protocols employed within the APCS. Examples of special ITI include computer attacks against the PTP (IEEE 1588) and Modbus protocols, CBCS controller firmware.

Step 6. Statistical estimation of CBCS protection consists in the estimation of CBCS protection based on the results of simulation and field modeling. CBCS protection assessment

is based on the verification of the correctness of control system operation per the levels of the reference model of open systems interactions.

Calculation of the statistical estimation indicators of CBCS protection is based on the ALARP principle and four established risk categories: from unacceptable to negligible [13, 14]. Accordingly, for each type of ITI and specified (obtained empirically or as part of simulation) the level of acceptable risk is selected along with the factor of relative risk scale spacing that allows formalizing the classification of ITI consequences as one of the four risk categories adopted in accordance with the ALARP principle. The values of the importance function are also selected individually for each ITI type.

Given the above, the final integral estimate of system risk can be calculated using the formula

$$R_{CBCS} = \frac{\sum_{j=0}^{3} k_j z_j w_j}{RS}, \quad (10)$$

where k_j is the number of ITI with the risk level j

z_j is the value of the significance function of the respective risk

w_j is the conditional weight of the respective level of risk

$RS = N \cdot k_3 \cdot w_3$, N is the number of the types of implemented ITI per threat model.

Additionally, the method suggests taking into consideration the designed level of CBCS protection that corresponds with its structural design and functional capabilities defined based on the additional indicators of Table 2.

Step. 7. Evaluation of the dynamics of the risk of CBCS protection disruption under ITI consists in the recalculation of the risk function, in which an additional coefficient is introduced that is based on the Weibull-Gnedenko distribution, which provides a dynamic CBCS protection estimate.

The dynamics of the risks of CBCS protection disruption under ITI is described with the fault (failure) rate function based on the Weibull-Gnedenko distribution function. The distribution density function is defined by formula (11), where α is the parameter of the shape of distribution that defines the nature of the risk dynamics throughout all life-cycle stages of the model (Figure 2), R_{BAS} is the parameter that defines the value of the basic risk coefficient of the distribution function

$$f(t_{LC}) = \frac{at^{a-1} e^{-\left(\frac{t_{LC}}{R_{BAS}}\right)^a}}{R_{BAS}}. \quad (11)$$

The values of the Weibull-Gnedenko distribution function are concentrated on the semiaxis from 0 to infinity. For experimental research of the dynamics of the risks of CBCS protection disruption under diverse and massive ITI and minimization of testing let us introduce a shift coefficient and use the hypothesis of the three-parameter Weibull-Gnedenko distribution, as well as the concept of risk function [4].

The function of the dynamics of assessment of the risk of CBCS protection disruption is developed out if the three-parameter distribution. The result of the transformations is given in (12). In order to obtain the function of the dynamics of the risk of CBCS protection disruption $R_{DYN}(t_{LC})$, the shift parameter must be taken into consideration that is defined based on the results of simulation and enables the correct shape of the risk dynamics for each type of CBCS. The function's behaviour is defined by (14)

Table 2. Additional indicators that characterize the designed protection of CBCS

Structural and functional characteristics of CBCS	Designed level of CBCS protection		
	High	Medium	Low
Based on the type of interaction between CBCS levels: - autonomous, no interlevel interaction, interaction is unidirectional (to the upper level) - bidirectional interaction.	3	2	-
Based on the input-output interface design: - local physical (electrical) interfaces - communication interfaces are used; they are physically separated from other network segments - common communication environment.	3	2	1
CBCS firmware replacement tools - absent - available, local connection required - remote control enabled.	3	2	1
Based on the inbuilt CBCS supervision and self-diagnostics tools: - absent - fault (failure) identification available - fault (failure) identification, operability recovery and functional redundancy facilities available.	3	2	1

$$R_{DYN}(t_{LC}) = \frac{a_k (t_{LC} - t_k)^{(a_k - 1)}}{\left(\frac{1}{R_{BAS}}\right)^{a_k}}, \quad k \in [1, 2, 3]. \qquad (12)$$

The values of index t_k are different for different lifecycle stages of the model (Figure 2). In [3], their generation is described in detail. The standard form of the risk dynamics curve will be defined by three model components, for each of which their own coefficient values are selected by means of simulation. The basic conditions are given in (13)

$$a_k : \{k = 1 : 0 < a_k < 1 \{k = 2 : a_k = 1,$$

$$t_k : \{k = 1 : t_k = 0 \{k = 2 : t_k = 0, \qquad (13)$$

$$R_{DYN}(t_{LC}) : \{a = 1 : R_{DYN}(t) = Const \{a > 1 : R_{DYN}(t) \uparrow. \qquad (14)$$

Graphical estimation and prediction of the evolution of the function of the dynamics of risk assessment of CBCS protection disruption under ITI based on the Weibull-Gnedenko distribution is shown in Figure 2.

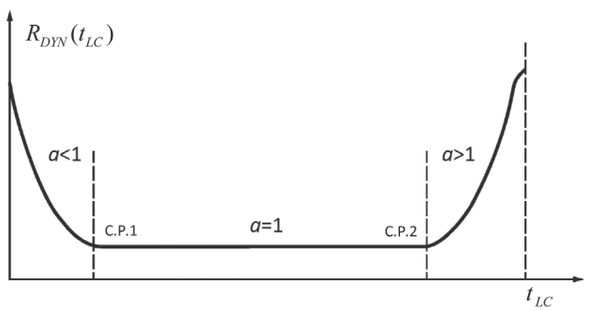


Figure 2. Standard form of the risk dynamics function

Risk dynamics (first stage, before characteristic point c.p.1) is defined by an increased level of risk due to the beginning of system operation and the following potential threats:

- fault (failure) of new version of software
- introduction of potential vulnerability in the new version of software and hardware
- insufficiently debugged information interaction between CBCS software and hardware facilities in the course of CC.

At the same time, the risk of CBCS protection disruption decreases over time due to CBCS software updates (patches), as well as improvement of the algorithms of control programs subject to the evolution of ITI threat model.

In Figure 2, risk dynamics $R_{DYN}(t_{LC})$ between two characteristic points c.p.1 and c.p.2 (second stage) are linear and correspond to the basic level of risk obtained by means of statistical estimation (10-12). This section corresponds to normal operation of a debugged system, when the patches are released regularly and the CBCS control algorithms have been debugged.

Within section after characteristic point c.p.2 (third stage) shows an even increase of the level of risk, which is both due to the possible decrease of hardware dependability, and accumulation of non-eliminated errors in software caused by zero-day ITI. Dynamic correction of risk $R_{DYN}(t_{LC})$ (formula (14) was generated is such a way as within the section between the points c.p.1 and c.p.2 (second stage) it has the value equal to one.

For a real CBCS, parameters $R_{DYN}(t_{LC})$ must be specified subject to the specificity of the system and the control cycle supervised by CADPS sensors. In the case of maximum possible compliance with the designed CBCS protection, the risk of protection disruption will be minimal. The adjustment of the value of the risk of protection disruption will be done based on the results of testbed experimental research under various ITI, as well as subject to risk dynamics over the control system lifecycle based on the risk function.

Conclusion. The proposed method of assessment of the protection of CBCS of critical information facilities enables numerical statistical and dynamic estimation of the risk of disruption of such system's protection under an intruder's ITI. The scientific novelty of the proposed method consists in the development of the model of CBCS operation under ITI based on augmented Petri nets and mathematics for the definition of the risk function of CBCS protection disruption using the Weibull-Gnedenko distribution.

The authors express their gratitude to Prof. I.B. Shubinsky for his assistance in the estimation of the integral risk of disruption of information protection of a control system.

References

[1]. Klimov SM, Kupin SV, Kupin DS. Models of malicious software and fault tolerance of information communication networks. Dependability 2017;4:36-43. DOI: 10.21683/1729-2640-2017-17-4

[2]. Collective of authors. «Umnye» sredy, «umnye» sistemy, «umnye» proizvodstva: seriya dokladov (seriya zelenykh knig) v ramkakh proekta «Promyshlennyy i tekhnologicheskiy forsayt Rossiyskoy Federatsii» ["Smart" environments, "smart" systems, "smart" plants: a series of reports (a series of green books) as part of the project Industrial and technological foresight of the Russian Federation]. Saint Petersburg: Center for Strategic Research North-West; 2012 [in Russian].

[3]. GOST R 50779.27-2017. Statistical methods. Weibull distribution. Data analysis [in Russian].

[4]. Kapur K, Lamberson L. Reliability in Engineering Design. Moscow: Mir; 1980.

[5]. Klimov SM, Astrakhov AV, Sychiov MP. Metodicheskie osnovy protivodeystvia kompiuternim atakam [Basic methods of computer attack response]. Moscow: Bauman MSTU; 2013 [in Russian].

[6]. Klimov SM, Astrakhov AV, Sychiov MP. Tekhologicheskiye osnovy protivodeystvia kompiuternim atakam

[Basic processes of computer attack response]. Moscow: Bauman MSTU; 2013 [in Russian].

[7]. Shubinsky IB. Nadiozhnie otkazoustoychivie informatsionnie sistemy. Metody sinteza [Dependable failsafe information systems. Synthesis methods]. Ulianovsk: Oblastnaya tipografia Pechatny dvor; 2016 [in Russian].

[8]. GOST R 56546-2015. Information protection. Vulnerabilities in information systems. The classification of vulnerabilities in information systems [in Russian].

[9]. Klimov SM, Kotyashev NN. Method of risk management for automated systems under conditions of cyber attacks. Dependability 2013;2:101-107 [in Russian].

[10]. Antonov SG, Klimov SM. Method for risk evaluation of functional instability of hardware and software systems under external information technology interference. Dependability 2017;17(1):32-39.

[11]. Klimov SM, Polovnikov AYu, Sergeev AP. A model of function-level fault tolerance of navigation signals provision processes in adverse conditions. Dependability 2017;17(2):41-47.

[12]. Klimov SM, Polikarpov SV, Fedchenko AV. Method of increasing fault tolerance of satellite communication networks under information technology interference. Dependability 2017:17(3):32-40.

[13]. Gapanovich VA, Shubinsky IB, Zamyshliaev AM. Risk assessment of a system with diverse elements. Dependability 2016;16(2):49-53.

[14]. Gapanovich VA, Rozenberg EN, Shubinsky IB. Some concepts of fail-safety and cyber protection of control systems. Dependability 2014;2:88-94 [in Russian].

About the authors

Sergey M. Klimov, Doctor of Engineering, Professor, Head of Division, 4-th Central Research and Design Institute of the Ministry of Defence of Russia, e-mail: klimov.serg2012@yandex.ru

Yuri V. Sosnovsky, Candidate of Engineering, Senior Lecturer, Department of Computer Engineering and Modeling, Physics and Technology Institute, V.I. Vernadsky Crimean Federal University, e-mail: yuri.sosnovskij@yandex.ru

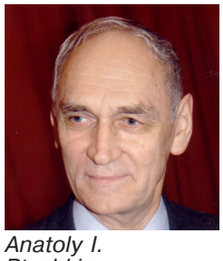
Received on 23.08.2018

Original article
Dependability, vol. 18 no.4, 2018 DOI: 10.21683/1729-2646-2018-18-4-45-50

Risk as one of the properties of decisions taken under uncertainty

Vladimir I. Zviagin, *Mozhaisky Military Space Academy, Saint Petersburg, Russia*
Anatoly I. Ptushkin, *Mozhaisky Military Space Academy, Saint Petersburg, Russia*
Alexey V. Trudov, *Mozhaisky Military Space Academy, Saint Petersburg, Russia*

Vladimir I. Zviagin



Anatoly I. Ptushkin

Alexey V. Trudov

Abstract. Aim. *The variety of activity types and the corresponding unfavorable outcomes has led to a dramatic diversity in terminology interpretations of the concepts related to risk, including ones in regulatory documents. This circumstance contradicts the primary purpose of the scientific and technical terminology standardization, which is the establishment of unambiguous and non-contradictory terminology in all types of documentation involved in the standardization activities or using the results of such activities. Given the above, this paper aims to estimate the conformity of the definitions of the concept of "risk" in the set of risk management standards with the requirements of the regulatory documents of the Russian standardization system and development of proposals regarding a new interpretation of this concept.* **Methods.** *The need for updating the existing definitions of risk and developing a definition complying with all the requirements of the regulatory documents of the Russian standardization system was based on the methods of terminological, logical-semantic and system analysis.* **Results.** *An analysis of compliance of the existing definitions of the term "risk" with the requirements of the Russian standardization system has been conducted and revealed that none of them fully meets such requirements, therefore the interpretation of the concept of "risk" requires a revision. The paper substantiates the interpretation of risk as one of the properties of the quality of a decision made in situations of uncertainty. This property characterizes the possibility and consequences of not achieving the goal of human activities in the situation of decision-making regarding the selection of further actions under uncertainty. Hence is given the following new definition of the term risk, i.e. one of the properties of the quality of a decision made in a situation of uncertainty that characterizes the possibility and consequences of not achieving the stated goals. The advantages of the proposed interpretation of risk over the existing definitions have been considered.* **Conclusions.** *The paper proposes and substantiates a new definition of the term "risk" that can be considered preferable over the existing versions. The proposed definition is based on most important concepts in terms of the theory and practice of management, i.e. "property", "quality", "decision", "situation", "goals" that are among the basic categories of human knowledge. This enables the use of both the existing quantitative characteristics of risk and the extension of the system of substantiated characteristics of risk, including those borrowed from the toolboxes of assessment of the manifestation rate of various objects' properties adopted in other domains of science. The authors show such special features of the proposed interpretation of risk as complexity, situation awareness and goal orientation. The complex and goal-oriented nature of risk prompts to consider actual capabilities to achieve the target characteristics of safety, performance, resource intensity and timeliness of reaching the specified goals of activities. The situational nature of risk as a property of a decision in a specific situation prompts the examination of the entirety of the associated contributing properties of the situation, i.e. the composition of the objects and subjects of human activity, as well as the conditions and circumstances that create specific relations between them. This approach significantly improves the precision of identification of the inventory and nature of the risk factors and therefore expands choices of risk management means and methods.*

Keywords: *term, risk, property, quality, situation, uncertainty of outcome, administrative decision, failure to achieve goal, characteristics, indicators.*

For citation: *Zviagin VI, Ptushkin AI, Trudov AV. Risk as one of the properties of decisions taken under uncertainty. Dependability 2018;4: 45-50. DOI: 10.21683/1729-2646-2018-18-4-45-50...*

Introduction

For a long time, the concept of risk has been associated with the possibility of some unfavorable (undesirable) outcome of a certain activity in the context of incomplete information about its further course. Such outcome could be damage or conflict of some kind (material or economic damage, damage to life or health of people and animals, environmental damage, political conflict, etc.). Risk assessment is mandatory for justifying nearly all important decisions.

When it comes to risk, the variety of activity types and the corresponding unfavorable outcomes has led to a dramatic diversity in terminology interpretations, including ones in regulatory documents (RD). This contradicts the primary purpose of the scientific and technical terminology standardization [1], i.e. the establishment of unambiguous and non-contradictory terminology in all types of documentation involved in the standardization activities or using the results of such activities.

Hence, one of the main tasks of scientific and technical terminology standardization is the analysis, identification and correction of the terminology shortcomings, most notably in national standards. The importance of this research area has long been emphasized in the works of prominent specialists in such knowledge-intensive, energy-intensive and risk-sensitive branches as nuclear and radiation safety [2], dependability and safety of structurally complex systems [3, 4], dependability and safety of pneumohydraulic components of space rocket systems [5], etc. The significance of this problem is most evidently expressed in [2]: "… our whole life largely depends on the clarity of regulatory documents… The clarity of terminology is the basis of both the scientific problems formulation and the regulatory laws adoption". The relevance of this task in terms of the research od substantial interpretation of such important concepts as dependability, safety and risk is being confirmed by a number of recent publications of a terminological nature, for example [6, 7, 8].

This paper aims to estimate the conformity of the definitions of the concept of "risk" in the set of risk management standards with the requirements of the regulatory documents of the Russian standardization system and to develop proposals regarding a new interpretation of this concept.

1. Subject and aim of the requirements of the Russian standardization system analysis

To make the following provisions more constructive and specific we shall narrow down the considered subject matter by restricting it to the following two conditions:

- instead of individual GOST standards the analysis will cover the set of existing "risk management" standards with practically identical subjects of standardization and application areas. This set currently includes over 25 standards! However, not all the standards in this set contain definitions of risk, so only those document where such definitions are given [10-24] are analyzed;

- the analysis is carried out in order to estimate the conformity of the terminological provisions in the considered set of standards with the requirements of the regulatory documents of the Russian standardization system. To achieve this goal, the following three questions should be answered.

Firstly, how justified is the above diversity of views on the concept of "risk"? The most logical explanation could be a significant dependence of the definitions on the characteristics of the activity types. Therefore, the first thing that needs to be clarified is whether the existing definitions of risk depend on the specifics of a particular activity.

Secondly, if the analysis of the first question reveals that the existing definitions of risk are activity invariant, then the next question arises: is there, among many existing definitions, one that best meets the requirements of the standardization system RD and therefore can be used (recommended) as a universal, generally accepted definition of risk?

And thirdly, if the answer to both previous questions is negative, the last question arises: what definition of risk can be proposed as acceptable in terms of conformity with the requirements of the standardization system RD?

The following subsections of the article are devoted to finding answers to these questions and developing the corresponding proposals.

2. Review of existing definitions of the term "risk" in the current standards

The analysis of the above set of documents [10-24] showed that there are mainly three definitions of risk:

- risk is the probability of causing harm to the life or health of citizens, property, the environment, etc. subject to the severity of this harm [10]. In [11], the definition is almost identical with the difference being that it proposes a qualitative measure of the possible severity of harm in addition to the quantitative measure;

- risk is the combination of the probability of hazardous event and its harm [12-21];

- risk is the effect of uncertainty on objectives [22-24].

The definitions from the science and technical publications where the risk is understood as a hazardous (undesirable) event, measure of danger, taking pot luck, loss of opportunity etc. can be added to these definitions of risk in RD. These and other definitions are followed by lengthy risk classifications by a variety of different attributes, for example, by hazard type (technology-related, natural, etc.), by field (financial, economic, ecological risks, etc.), by consequence scale (high, moderate, low, critical, catastrophic, etc.). Unfortunately, neither these definitions and classifications nor the names, functions and properties of risks listed in them clarify the essence of the concept itself leaving the question open.

The answer to the first of the questions follows from the above definitions of risk, namely *the existing definitions of risks in RD and scientific and technical literature are*

rather general and are not related to the characteristics of a particular activity. Therefore the existing diversity of these definitions in RD can hardly be considered useful for theory and practice.

3. Analysis of compliance of the existing definitions of the terms "risk" with the requirements of the Russian standardization system

The primary, guiding principle of the Russian standardization system RD [25-29] is the unambiguousness of the requirements included in the standardization documents [26]. To implement this principle, a set of basic requirements of this system should be met. Table 1 shows a summary of those requirements.

Table 1. Requirements of the standardization system regulatory documents for the generation of terms and their definitions

№	Requirement content
1	One designation (i.e. term, symbol or name) shall correspond to one concept, and only one concept shall correspond to one designation
2	One term shall not be used for many concepts and many terms shall not be used for one concept
3	Terminological entries in closely related standards shall not be contradictive
4	A term shall show the limiting characteristics of the concept expressed
5	A term shall maintain the usual form of expression established in the speech community
6	A term shall correspond to the morphological, morphosyntactic and phonological norms of the language
7	Native language shall be prioritized
8	A term definition shall be a single phrase defining the concept and, if possible, reflecting its position in the system of concepts

Note that the requirements 4 to 6 in Table 1 apply not only to terms, but also to their definitions. The first thing to observe when analyzing the listed requirements is that the aforementioned definitions of risk have different concepts for the same term, which contradicts requirements 1 and 2. Moreover, there is a contradiction between the first two and the third interpretation of risk. Indeed, in the former case the risk appears as a measure, and in the latter as a certain effect of uncertainty, which can be interpreted as you please: numerically, qualitatively, with verbal constructs, etc., which does not comply with requirement 3.

Speaking of each definition of risk separately, the third interpretation is the most arguable, since although it is present in harmonized terminological standards [22, 23], it does not meet the requirements 4 to 6.

The last requirement in the table can be best met by using intensional definitions [29]. The basic part of the intensional definition defines the superordinate concept that includes the concept in question, and the second part specifies the limiting characteristics that differentiate this concept from its peer concepts. At this point the shortcomings of the most common definitions of the concept of "risk" that are the first two interpretations should be mentioned. Unfortunately, they don't use intensional definition with the corresponding basic part and limiting characteristics, but simply identify risk with its measure (as already mentioned above), i.e. with one of the possible numerical characteristics, which is a confusion of different semantic categories and contradicts requirement 8. In addition, replacing the semantic interpretation with a numerical characteristic (indicator) contradicts other standards of the "risk management" set (see, for example, [20, 21]), that mention such indicators as risk index and risk severity index that are similar in meaning to the mentioned risk interpretations, but describe the level of risk, not the risk itself.

Thus, *the analysis of existing risk definitions has shown that they are far from fully complying with the basic principles and requirements of the Russian standardization system. Hence the conclusion that the concept of "risk" requires revising*, and that is what the next subsection of this article is devoted to.

3. Proposed interpretation of the concept of "risk" that complies with the requirements of the Russian standardization system

First of all, it should be noted that in the existing scientific and technical publications, much less in the RD, no substantiation of the above risk definitions could be found. In almost all cases, they are simply postulated, often with numerous notes supplementing the proposed definition with its possible interpretations (see, for example, [22-24]). The undesirability of this approach is illustrated above.

The substantiation of the interpretation of the term "risk" can be based on terminological, logical-semantic and system analysis methods and is as follows.

Human experience shows that the concept of "risk", whatever the understanding, is nearly always associated with the situation of decision-making under uncertainty. This situation is about making a choice among a set of alternatives when the information on the possibility and nature of the decision consequences is insufficient. This situation is most typical for management decisions that are made with lacking initial information. Such close relation between risk and management decisions allows us using some findings of the management theory widely covered in management-related literature.

One of the central concepts in this theory is the quality of managerial decisions, that by analogy with the widely

accepted interpretation of product quality in [32], means (see, for example, [30, 31]) a set of decision properties (characteristics, parameters) that fulfil a certain function in the management process and satisfy a specific consumer. The indispensable attributes of decision-making under uncertainty are the possibility and consequences of not achieving this decision's goals.

What remains to say is that management decision quality should be measured, it is therefore impractical to conceptually deviate from the recommendations of qualimetry, a scientific discipline that concerns the methods and problems of quantification of the quality of any object, according to which quality is a combination of properties of an object that a person deals with in practical activities.

As a result, it seems logical and reasonable to include the decision riskiness (risk) into the set of such properties. It is the property that characterizes the possibility and consequences of not achieving the goals of a human activity when making a decision under uncertainty. Then the following final definition of the term "risk" can be proposed:

Risk (riskiness) is one of the properties of the quality of a decision made in a situation of uncertainty that characterizes the possibility and consequences of not achieving the stated goals.

At the same time, when conducting a scenario analysis of a situation, the entire potential semantics of the possible outcomes (positive, negative, neutral, etc.) should be taken into consideration.

Briefly summarizing, the main system features of the proposed definition are as follows. Firstly, it interprets risk not as a measure, but as a certain attribute of a decision being made, namely, its property. Secondly, risk as a property is a component of the quality of this decision being a kind of management act. Thirdly, risk as a decision property in goal-oriented processes is complex. Indeed, any human activity is associated with the need to achieve at least four goals: ensuring the required safety, performance, resource-intensiveness and timeliness characteristics. The multi-objective nature of this activity requires risk to be considered as a complex property, which includes properties that characterize the possibility and consequences of actual goal achievement values not meeting the required values. Table 2 shows examples of not achieving these goals in technical areas.

Finally, the fourth feature of the given risk definition is its situational nature, dependency on the decision-making situation, i.e. the composition of objects and subjects of human activity, the conditions and circumstances that create a certain relationship between them.

It should be noted that decision-making should be preceded by a set of preparatory procedures, including, for example, forecasting the possible consequences of catastrophic climatic and natural phenomena. Based on the results of such procedures, a decision should be made on the nature and sequence of further actions to achieve the goal.

Table 2. Examples of not achieving activity goals

Goals	Examples
Provision of safety	Death of personnel or citizen, harm to health of personnel or citizens, equipment or property of citizens, the environment
Performance	Failure to achieve the required values of product quality indicator, failure to perform the task
Achievement of the required result with the allocated resources	Excessive consumption of allocated resources (material, financial etc.)
Timely achievement of the required result	Delays in completion of works at various stages of achieving the result

Conclusions

The above substantiation and features of the proposed interpretation of risk suggest the following advantages over the existing definitions.

1. First of all, this definition is free from the shortcomings of the existing definitions and meets all the requirements of the Russian standardization system. Therefore, the proposed definition of risk can be considered preferable over the existing ones, that, as the article shows, do not meet the requirements of this system.

2. The proposed definition is based on the most important concepts in terms of the theory and practice of management, i.e. "decision", "property", "quality", "goal", "situation" that are among the basic categories of human knowledge and largely define the differentiation of sciences. Risk as a property, as an aspect of quality has a certain intensity, i.e. it can be "major" or "minor", "high" or "low", "acceptable" or "unacceptable", etc. This enables the extension of the system of substantiated characteristics of risk, both qualitative and quantitative. For example, interpreting risk as of one the properties of quality enables the use of methodology for assessing the manifestation rate of various objects' properties adopted in other domains of science, including qualimetry, dependability theory, game theory, operations research, etc. For instance, following a well-developed conceptual framework of such important technical property as dependability [33], the most well-established and proven terms and concepts can be borrowed from this domain. In particular, risk as a complex property can be differentiated into a number of its particular properties. At the same time, risk as a property of a human decision is objective in nature, although its estimation may also have signs of subjectivity, and rightly so.

3. The proposed interpretation of risk does not cancel, but allows the use of the existing risk characteristics that can be found in literature, such as probability of a risk realization scenario, extent of damage (harm), their combination, risk level, risk index [20, 21], etc. Numerous classifications of risk by type of activity and other characteristics, consider-

ing the stages and phases of risk requirements development, risk analysis and risk management, etc., i.e. the whole set of tools of research of risk as a full-fledged scientific category, remain in force.

4. The "decision situation" featured in the risk definition prompts the examination of the entirety of the associated contributing properties of the situation when choosing and considering risk characteristics. Therefore, the analysis of all undesirable event scenarios (scenario analysis) of this and other potential situations that may arise from implementing the decision should be an essential component of risk assessment. This approach significantly increases the precision of identifying the inventory and nature of the risk factors and therefore extends the options of risk management means and methods:

5. This goal-oriented interpretation of risk directly implies a number of important requirements to the management activities organization including:

- any management decision should be goal-oriented, i.e. it should include the estimation of the characteristics of the possibility and consequences of not achieving the decision goals;
- generation of decision requires combined consideration of mutual influence of risk components related to all the decision goals;
- to improve the quality of coordination of the interests of the parties involved in the decision implementation and the use of allocated resources, the whole range of institutional relationships among them should be taken into account;
- to improve the precision of identification of the inventory and nature of the risk factors and to extend the options of risk management means and methods, the whole set of situational characteristics should be considered when developing alternative solutions.

In conclusion, it should be mentioned that the term "risk" is widely used in various domains and therefore requires close attention of all stakeholders. The authors are convinced that the proposed interpretation of risk as a property of the quality of a decision made in situation of uncertainty is productive in terms of the risk management theory development. However, they are well aware that the proposed definition and its justification are not flawless, and therefore their constructive criticism as part of the corresponding discussion could be useful.

References

[1]. R 50.1.075-2011. Development of Standards on Terms and Definitions. Approved and brought into force by order of the Federal Agency for Technical Regulation and Metrology no. 35-st. dated March 28, 2011 [in Russian].

[2]. Gordon BG. Ob ispolzovanii ponyatiya riska v razlichnykh otraslyakh promyshlennosti [On the use of the concept of risk in various industries]. Vestnik Gosatomnadzora Rossii 2003;1:3-7 [in Russian].

[3]. Riabinin IA. Nadezhnost i bezopasnost strukturno-slozhnykh sistem [Dependability and safety of structurally complex systems]. Saint Petersburg: Saint Petersburg State University Publishing; 2007 [in Russian].

[4]. Shubinsky IB. Strukturnaya nadezhnost informatsionnykh sistem: metody analiza [Structural dependability of information systems: methods of analysis]. Ulianovsk: Pechatny dvor; 2012 [in Russian].

[5]. Golikov IO, Grankin BK. O sostoyanii normativnogo obespecheniya slozhnykh tekhnicheskikh kompleksov [On the state of the art of the regulatory support of complex technical systems]. Standards and quality 2016;5(947):76-80 [in Russian].

[6]. Golikov IO, Grankin BK, Zviagin VI. Obuchenie kachestvu po standartam [Teaching standard-based quality]. Standards and quality 2017;2(956):29-33 [in Russian].

[7]. Pokhabov Yu.P. On the definition of the term «dependability». Dependability 2017;1:4-10.

[8]. Strukov AV. Analiz mezhdunarodnykh i rossiyskikh standartov v oblasti nadezhnosti, riska i bezopasnosti [Analysis of the international and Russian standards in the area of dependability, risk and safety], <https://szma.com/standarts_analysis.pdf>; [accessed on 07.06.2018].

[9]. GOST 1.1-2002. Interstate system for standardization – Terms and definitions. Moscow: Izdatelstvo standartov; 2003 [in Russian].

[10]. Federal Law no. 184-FZ dated December 27, 2002 On technical regulation (as revised on July 29, 2017).

[11]. GOST R ISO 11231-2013 Risk management. Probabilistic risk assessment through the example of space systems. Moscow: Standartinform; 2014 [in Russian].

[12]. GOST R ISO 12100-1-2007. Safety of machinery. General principles of design. Part 1. Basic terms, methodology. Moscow: Standartinform; 2008 [in Russian].

[13]. GOST R ISO 17666-2006. Risk management. Space systems. Moscow: Standartinform; 2006 [in Russian].

[14]. GOST R ISO-IEC 31010-2011. Risk management. Risk assessment methods. Moscow: Standartinform; 2012 [in Russian].

[15]. GOST R 51901-2002. Risk management. Risk analysis of technological systems. Moscow: Izdatelstvo standartov; 2002 [in Russian].

[16]. GOST R 51901.4-2005. Risk management. Application guidelines for projects. Moscow: Standartinform; 2005 [in Russian].

[17]. GOST R 51901.11-2005. Risk management. Hazard and operability studies. Application guide. Moscow: Standartinform; 2006 [in Russian].

[18]. GOST R ISO-IEC 16085-2007. Risk management. Application for system and software life cycle processes. Moscow: Standartinform; 2008 [in Russian].

[19]. GOST R 56255-2014. Terms and definitions in the field of life and health safety. Moscow: Standartinform, 2015 [in Russian].

[20]. GOST R 54144-2010. Risk management. Implementation guide for organizational security measures and risk assessment. Incident identification. Moscow: Standartinform; 2012 [in Russian].

[21]. GOST R 54145-2010. Risk management. Implementation guide for organizational security measures and risk assessment. General methodology. Moscow: Standartinform; 2012 [in Russian].

[22]. GOST R 51897-2011. Risk management. Terms and definition. Moscow: Standartinform; 2012 [in Russian].

[23]. GOST R ISO 9000-2015. Quality management systems. Fundamentals and vocabulary. Moscow: Standartinform; 2015 [in Russian].

[24]. GOST R ISO 31000-2010. Risk management. Principles and guidelines. Moscow: Standartinform; 2012 [in Russian].

[25]. Federal Law no. 162-FZ dated June 29, 2015 On standardization in the Russian Federation (as revised on July 3, 2016) [in Russian].

[26]. GOST R 1.0-2012 Standardization in the Russian Federation. Basic provisions (as revised on July 1, 2014). Moscow: Standartinform; 2013 [in Russian].

[27]. GOST R 1.2-2016. Standardization in Russian Federation. National standards of Russian Federation. Instructions for development, taking over, revision, correction, suspension and cancellation. Moscow: Standartinform, 2016 [in Russian].

[28]. GOST R 1.5-2012. Standardization in Russian Federation. National standards. Rules of structure, drafting, presentation and indication. Moscow: Standartinform, 2013 [in Russina].

[29]. GOST R ISO 10241-1-2013. Terminological entries in standards. Part 1. General requirements and examples of presentation. Moscow: Standartinform; 2015 [in Russian].

[30]. Porshnev AG, Razu ML, Tikhomirova AV. Menedzhment: teoriya i praktika v Rossii: Uchebnik [Management: theory and practice in Russia: Textbook]. Moscow: ID FBK- PRESS; 2003 [in Russian].

[31]. Prokhorov YuK, Frolov VV. Upravlencheskie resheniya: Uchebnoe posobie [Administrative decisions: Teaching aid]. Saint Petersburg: SPbGU ITMO; 2011 [in Russian].

[32]. GOST 15467-79. Product quality control. Basic concepts, terms and definitions. Moscow: Standartinform; 2009 [in Russian].

[33]. GOST 27.002-2015. Dependability in technics. Terms and definitions. Moscow: Standartinform; 2016 [in Russian].

About the authors

Vladimir I. Zviagin, Candidate of Engineering, Professor, Professor of the Department of Organization of armament operation and technical support, Mozhaisky Military Space Academy, Saint Petersburg, Russia, e-mail: v.zvyagin@mail.ru

Anatoly I. Ptushkin, Candidate of Engineering, Professor, Professor of the Department of Organization of armament operation and technical support, Mozhaisky Military Space Academy, Saint Petersburg, Russia, e-mail: anatoly.ptushkin2011@yandex.ru

Alexey V. Trudov, Candidate of Engineering, Professor, Associate Professor of the Department of Organization of armament operation and technical support, Mozhaisky Military Space Academy, Saint Petersburg, Russia, e-mail: e-mail: trudovalex2016@yandex.ru

Received on 28.06.2018

Original article
Dependability, vol. 18 no.4, 2018 DOI: 10.21683/1729-2646-2018-18-4-51-55

Ensuring an efficient transportation infrastructure security system by means of solutions that enable detection of intrusions into protected areas

Natalia A. Kuzmina, *Far Eastern State Transport University, Khabarovsk, Russia*

Natalia A. Kuzmina

Abstract. *Due to the nature of its operations, the transportation industry in itself is a potential source of danger. In case of unlawful aggressive intrusions the danger becomes real and fraught with grave consequences. The statistics of the last 10 to 15 years show that 50 to 70% of accomplished terrorist attacks were associated with transportation. Individual measures cannot ensure transportation security. The problem must be approached comprehensively and systemically. Transportation security greatly contributes to the national security of the Russian Federation. The Federal Law of February 9, 2007 no. 16-FZ On transportation security, for the first time in Russian practice, raised the question of securing the entire transportation industry of the Russian Federation, established the legal foundations of the activities related to the protection of transportation infrastructure and vehicles against acts of unlawful interference, including those of terrorist nature. For the first time, a single systemic approach to anti-terrorist protection is provided for all means of transportation. The transportation industry is quite vulnerable to terrorist attacks. We are talking about vehicles, transportation lines, stations, vehicles carrying dangerous loads. The vulnerability of transportation is due to the possibility of damage to signalling, automation and communication assets, whose protection is complicated due to the scale and extent of Russia's railways. Despite the problems and objective difficulties related to the legislation in the area of transportation security, the workers of the Russian railway industry make their best effort to ensure protection of transportation infrastructure and vehicles against acts of unlawful interference. Promptly reacting to other challenges and threats, they ensure reliable operation of the transportation industry, thus preserving the peace and safety of our citizens. This paper examines matters related to ensuring efficient safety of transportation infrastructure. A significant emphasis is placed on the systems that enable detection of intrusions into protected areas of a facility.*

Keywords: *transportation security, act of unlawful interference, transportation infrastructure facility, security system, facility protection, intruder, sensors, annunciator.*

For citation: *Ensuring an efficient transportation infrastructure safety system by means of solutions that enable detection of intrusions into protected areas. Dependability 2018;4: 51-55. DOI: 10.21683/1729-2646-2018-18-4-51-55*

Despite the measures taken by transportation infrastructure managers and carriers that aim to improve the protection of the transportation industry, the threat of acts of unlawful interference (AUI) remains. Statistics show that, unfortunately, it is not yet possible to completely eliminate the possibility of an act of unlawful interference. In this regard, the priority task is to hinder AUI attempts against transportation infrastructure facilities and vehicles as well as to disrupt the intruders' plans.

This is only possible if the government and the society deliver a consolidated approach to the security issues [1].

The Federal Law no. 16-FZ On transportation security makes it compulsory to protect transportation infrastructure facilities against AUIs [2].

Protection, as a rule, involves the availability of technical means of protection that correspond with the facility's category.

This approach does not include an efficiency assessment of the measures taken and the ability of the whole system to meet real threats. There is no guarantee that in case of an emergency situation the response will be prompt and correct.

What should be an effective security system for the facility and how to protect it? This issue is relevant today for many transportation infrastructure facilities, which face the difficult task of ensuring transportation security.

At the beginning of 2014, the Federal Law no. 15-FZ On Amendments to Certain Legislative Acts of the Russian Federation on Transportation security came into force. With this document the gaps of the Federal Law no. 16-FZ On transportation security issued in 2007 were eliminated. This basic law had an extremely broad interpretation, and did not contain specific requirements to the forces and means of ensuring transportation security, and in some areas only formalized the process [3].

Under a comprehensive approach to ensuring the security of transportation infrastructure facilities, the reasons that enabled a terrorist attack should be considered as symptoms of unsatisfactory operation of such facilities, whose improvement must be the focus of the efforts, application of human and material resources.

According to the theory of the multiplicity of causes of emergency situations proposed by D. Peterson, Professor of University of Colorado (US), it is possible both to predict the possibility and identify the circumstances of their occurrence. Consequently, security, in transportation included, not only can, but should be controlled as any other part of the transportation system. Transportation security should be one of the inherent, daily functions of the chief executive officers and managers in various positions (among such functions could be cost reduction, enabling the required volume of freight and passenger traffic).

When solving problems of ensuring security of transportation infrastructure facilities, systems that allow detecting intrusions into protected areas play a major role. For this purpose different system control panels are used that notify about intrusions into protected areas with an indication of the place and time of violation of the area boundary. These systems ensure data collection from sensors that monitor the area of possible intrusions, detecting the fact of unauthorized entry and transmission of alarm to the system control panel. The system control panels are often integrated with fire warning systems and have practically the same structure.

The area protection system includes centralized control equipment, input control equipment (ICE) or panels (ICP) and control sensors of the protected area. A computer or special control panel that fulfills its function in some systems enables the centralized management of the system. ICE supplies power to loops with the associated sensors in the area, receives and analyzes the messages transmitted by the sensors, generates and transmits the alarm to the centralized control station, controls the warning devices and other security systems. In small facilities the ICE can control all systems without data transmission to the centralized control station.

Various annunciator sensors that differ in type, operating principle and functionality are employed in system control panels to monitor the areas of potential intrusion. Among such sensors are magnetic, vibration, photoelectric, microwave, ultrasonic sensors, loops, glass break sensors, motion sensors.

The most simple, cheap and widely used sensors are *magnetic contact sensors* installed on windows and doors of the protected facilities. These sensors include a magnetically controlled contact (seal switch) installed on the moving part of the window and a permanent magnet installed on the window or door frame.

An example of such sensor is the IO-102-14 (SMK-14, seal switch) compact removable magnetic contact security annunciator used for doors and windows protection (Figure 1, a).

Glass break sensors (Figure 1, b) react to the sound of breaking glass and are designed for the windows of protected facilities. The principle of sensor operation is based on the spectrum analysis of the detected noise and its comparison with the reference sound signals recorded in the sensor memory. The most advanced sensor types sound a warning signal in two cases: glass being hit and sound of breaking glass.

Vibration sensors (Figure 1, c) detect vibrations and shocks associated with attempts to destroy the protected facility. The operation principle of such sensors is based on the piezoelectric effect or electromagnetic induction for conversion of vibration signals into analyzed electric signals.

Photoelectric infrared (beam) sensors (Figure 1, d, e) are used for protection of corridors, flights of stairs, gates and entrances, as well as for the facility perimeter protection.

At the core of sensor operation is a barrier of modulated infrared beams created by special emitters for the purpose of protection of a facility's secured area. When an intruder crosses the barrier's sensitivity zone, an alarm is triggered.

Loops of security alarm systems are bands of aluminum foil glued to components of protected structures (glass, door etc.) and included in the security alarm circuit. If a structure with a band is destroyed, the security alarm circuit is broken and an alarm is triggered.

Motion sensors allow detecting movement in protected areas based on the changes in the intensity of infrared radiation when heat-emitting objects move within the sensor's sensitivity zone. As an object moves, infrared radiation reflects from different segments of the optic system that causes the generation of a number of impulse signals detected by the

a *b* *c* *d* *e*

Figure 1. a) IO-102-14 (SMK-14, seal switch) magnetic contact annunciator; b) IO329-3 Arfa glass break sensor; c) Shorokh-2 (IO-313-5/1) surface vibration security annunciator; d) NR110QS four-beam infrared barrier sensor; e) application scheme of radial sensors

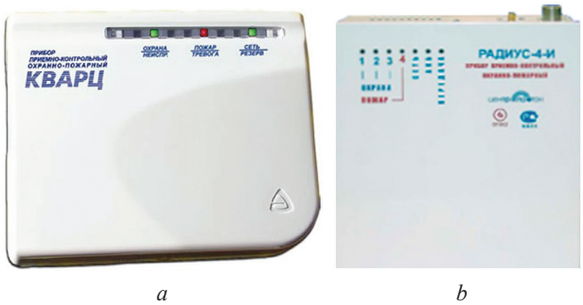
electronic sensor system. To protect against false triggering, today's motion sensors use microprocessors that process detected signals. Two types of infrared passive annunciators of the Foton family are presented as an example of motion sensors in Figure 2.

Along with the above types of annunciators, radio-wave, ultrasonic, micro-wave, capacitance-type sensors can be used in a security alarm system, whose operating principles are based on the analysis of signals reflected from an object or on the changes in the area and capacity of the protected facility. However, these sensors are much less widely used, since they are very sensitive to the environmental changes, type of the protected facility and various destabilizing factors.

Alarm annunciators transmit information to another mandatory structural component of any security system, i.e. ICE or ICP in case of smaller facilities. The function of the ICE is hardware performance verification, collection and analysis of the information received from annunciators, alarm transmission to the security panel, management of light, sound and fire warning signalling, as well as management of other technical equipment and systems of facility protection. The Quartz, Radius, Signal, Rif devices are examples of input control, security and fire equipment.

Quartz (Figure 3, a) is a fairly simple ICE that enables power supply and supervision of one loop with connected security and fire annunciators.

This device also controls emergency voice alarm communication systems and sends warning signals to the central monitoring panel.



a *b*

Figure 3. Input control equipment: a) Quartz; b) Radius-4I

The Radius-4I and Radius-6I input control, security and fire alarm equipment (Figure 3, b) used in the railway industry can be applied both for centralized and autonomous protection of facilities against unauthorized access and fires.

Radius-4I allows monitoring 4 loops; Radius-6I allows monitoring 6 loops with associated security and fire annunciators. After receiving a signal from an annunciator, the device sends a warning signal via communication channels as well as alerts of protected area intrusion or fire by means of light and sound signals. Both devices indicate the loop status using LEDs. The operating temperature range of the device is -10°C to +50°C, which allows it to be used widely in the protection of various facilities.

a *b*

Figure 2. Motion sensors: a) Foton-9 infrared passive alarm sensor; b) Foton-21 electrooptical ceiling-mounted security annunciator

Protection of geographically distributed fixed facilities involves the use of radio communication-based security systems that use a radio channel for transmission of intrusion and fire alerts. Rif String-202 is an example of such systems. This system includes a base station located in the security center, a monitoring panel and a computer, as well as input control equipment with transmitting devices installed in the facilities. One control panel allows controlling up to 600 10 MW transmitting devices (installed in the facility). Depending on the local topography, the system's range of communication varies from 25 to 50 kilometers or more with the help of ultra-narrow band communication channels and noiseless coding. The system uses digital filtering, simultaneous and parallel processing of all signals received via different communication channels from the controlled facilities. The input control equipment transmits each new message over a new frequency randomly selected from 1024 preprogrammed communication frequencies. Equipment operability is verified with each facility transmitter sending a pilot signals to the base station once per minute. Rif String-202 operation does not require authorization documents due to the use of licensed frequencies and low-power transmitters.

For total control of large areas, radio communication-based security systems can also be used that collect data from radio warning sensors installed around the facility area. In this case sensors are installed around the whole territory of the protected facility in such a way as to make sure the distance between them does not exceed the sensor operation radius, which ensures foreign object detection and does not create so-called dead zones. When a person or a foreign object enter a sensor's sensitivity zone, the sensor identifies it as an emergency situation and sends a warning signal to the system control panel via the radio channel.

In many cases, it is sufficient to control the perimeter to ensure security and completely eliminate the possibility of an unauthorized entry into the territory. Such

systems are called perimeter security systems. They are indispensable for facilities taking up large areas, such as airfields, warehouses, transport and logistic hubs, as well as extended facilities, such as railways and motorways, pipe lines etc.

Perimeter security systems allow detecting trespassing of protected areas and sending a warning signal much earlier than an intruder can reach important facilities located around such area. These systems are often structurally related with the physical fence of a protected facility, therefore, their operation directly depends on the physical parameters of the fence, presence of vibrations, materials used, as well as the equipment locations, quality of installation and a number of other factors.

To ensure reliable facility protection, perimeter security systems shall meet specified requirements, such as:

- total control of the whole perimeter of the protected facility and absence of "dead zones"
- high sensitivity of intruder detection
- low probability of false triggering
- dependable operation in terms of electromagnetic interference from active equipment and industrial facilities located nearby
- dependable operation in various climatic conditions.

Currently, beam, radio wave and capacitance-type perimeter security systems are most often used to ensure facilities security. These systems have high efficiency of detection and allow reliable protection of an area. However, beam-based systems, such as infrared barriers can effectively control only straight sections of the perimeter. Radio wave-based systems are sensitive to surface geometry and work poorly if there are trees or bushes in the protected area, which is typical for railways and motorways. Capacitance-type systems require the presence of a physical fence around the facility, since they detect the electric field and capacity changes as the intruder approaches or touches such fence.

Wire and radio wave systems can also be used to protect extended facilities. Such systems consist of two parallel feeders installed along the protected area. One of the feeders serves as a receiving antenna, and the other is a transmitting antenna. If a foreign object gets within the feeder's operating range, field distortion ensues, which modifies the parameters of the received signal that is constantly picked up by the respective equipment.

The radio communication-based security systems under consideration are normally integrated with video surveillance systems, since potential intruders must be identified.

Naturally, all decision-making in a security system relies on humans. They monitor the status of technical assets, receive warning signals, control the response actions. Despite a clear action plan, strict compliance with the transportation safety requirements as regards railway infrastructure and rolling stock, in non-typical situations the natural human factor may come into action due to fatigue, confusion and even negligence. Then, an error or delay in responding to threats can cause loss of human life [4].

Over the last few years, a number of terrorist attacks were carried out in transportation facilities or using vehicles. Both Russian and international statistics are clear. Thus, ensuring the security of transportation infrastructure facilities and vehicles against intentional acts of unlawful interference in the form of terrorist attacks and sabotage has become a pressing problem around the world [5].

In 2017, 670 acts of unlawful interference were registered within the Russian railway system. Among them are cases of terrorist nature, including 68 reports of threats of terrorist attacks. 142 foreign objects were uncovered on tracks. 79 and 269 cases were registered of railway tracks and signalling systems dismantling, respectively.

Also in 2017, 5 cases of explosive substances and 665 cases of unattended suspicious objects were detected in railway facilities. 20 units of firearms, 1292 bullets of various calibers and 49 explosive objects were confiscated [6].

Meanwhile, the experience of the US, Canada and a number of European countries shows that the improvement of transportation security is a major concern not only for the Government. The problem is very urgent and public organizations are actively involved in it. For example, in the US hundreds of private companies and firms managed by non-governmental organizations regularly allocate significant shares of their annual budgets to research and development of solutions in the area of transportation security. The above observations show that both the concern of radical and urgent improvement of transportation security and allocation of financial, organizational and human resources should not be the burden of the Government alone. Additionally, even the best designed transportation security system cannot function effectively without an all-around support of the civil society.

Summarizing, it should be reminded that one of the primary tasks of a transportation infrastructure company or operator is to ensure human security. Security is one of the key conditions for the development of an individual, a society or a Government. This always requires people to be competent as regards the threats and methods of protection.

Application of physical protection assets in combination with organizational measures and actions of transportation security units is the primary factor of detection and response to acts of unlawful interference against property, freight and individuals in railway facilities.

An efficient security management system is like a good soldier who is responsible for the accurate and timely execution of assigned duties and tasks both in peace and war time. It also creates a comprehensive vision of the situation and minimizes the probability of a mistake [7].

References

[1]. Starovoytov AS. Zashchishchennost obiektov transporta neobkhodimo podderzhivat [Security of transportation facilities must be maintained]. Transportnaya i tekhnologicheskaya bezopasnost 2017;2:100-101 [in Russian].

[2]. Federal Law of 09.02.2007 no. 16-FZ On transportation security [in Russian].

[3]. Federal Law of 03.02.2014 no. 15-FZ On amendments to a number of legislative acts of the Russian Federation concerning transportation security [in Russian].

[4]. Order of the Government of the Russian Federation of 26.04.2017 no. 495 On the approval of transportation security requirements, including the requirements for antiterrorist security of facilities (territories) subject to security levels for various categories of transportation infrastructure facilities and railway vehicles [in Russian].

[5]. Kuzmina N, Odudenko T. Ensuring transportation security in the transport infrastructure and means of railway transport facilities. In: Proceedings of the XV International Academic Congress Fundamental and Applied Studies in the Modern World. Oxford (United Kingdom): Oxford University Press»; 2016.

[6]. <www.roszeldor.ru.>

[7]. Protsess evoliutsii ne ostanovit [The evolutionary process is unstoppable]. Transportnaya i tekhnologicheskaya bezopasnost 2017;2:34 [in Russian].

About the author

Natalia A. Kuzmina, Candidate of Pedagogy, Associate Professor of the Department of Transportation Management and Transportation Security, Far Eastern State Transport University, Khabarovsk, Russia, e-mail: kuzminaprepodavatel@mail.ru

Received on 10.01.2018

Gnedenko e-Forum

International Group on Reliability

http://Gnedenko-Forum.org/

Dear colleagues!

In 2005 the informal Association of Experts in Reliability, Applied Probability and Statistics (I.G.O.R.) was established with its own Internet website GNEDENKO FORUM. The site has been named after the outstanding mathematician Boris Vladimirovich Gnedenko (1912-1995). The Forum's purpose is an improvement of personal and professional contacts between experts in the mathematical statistics, probability theory and their important branches, such as reliability theory and quality control, the theory of mass service, storekeeping theory, etc.

Since January 2006, the Forum has published a quarterly international electronic magazine

"Reliability: Theory and Applications".

The magazine is registered with the Library of Congress in the USA (ISSN 1932-2321). All rights reserved for authors so that articles can be freely published in any other publications or presented at conferences.

Join Gnedenko Forum!

Welcome!

More than 500 experts from 44 countries worldwide have already joined us!

To join the Forum, send a photo and a short CV to the following address:

Alexander Bochkov, PhD

a.bochkov@gmail.com

Membership is free.

SUBSCRIBER APPLICATION FOR DEPENDABILITY JOURNAL

Please subscribe us for 20___
from No. _____ to No._____ number of copies _____

Company name	
Name, job title of company head	
Phone/fax, e-mail of company head	
Mail address (address, postcode, country)	
Legal address (address, postcode, country)	
VAT	
Account	
Bank	
Account number	
S.W.I.F.T.	
Contact person: Name, job title	
Phone/fax, e-mail	

Publisher details: Dependability Journal Ltd.
Address of the editorial office: office 209, bldg 1, 27 Nizhegorodskaya Str., Moscow 109029, Russia Phone/fax: 007 (495) 967-77-02, e-mail: evgenya.patrikeeva@yandex.ru
VAT 7709868505 Account 890-0055-006
Account No. 40702810100430000017
Account No. 30101810100000000787

Address of delivery:

To whom:_____

Where:_____

_____

To subscribe for Dependability journal, please fill in the application form and send it by fax or email.
In case of any questions related to subscription, please contact us.
Cost of year subscription is 4180 rubles, including 18 per cent VAT.
The journal is published four times a year.

THE JOURNAL IS PUBLISHED WITH PARTICIPATION AND SUPPORT

OF JOINT-STOCK COMPANY RESEARCH & DESIGN INSTITUTE
FOR INFORMATION TECHNOLOGY, SIGNALLING AND TELECOMMUNICATIONS
ON RAILWAY TRANSPORT (JSC NIIAS)

JSC NIIAS is RZD's leading company in the field of development of train control and safety systems, traffic management systems, GIS support technology, railway fleet and infrastructure monitoring systems

Mission:

transportation

- ❑ efficiency,
- ❑ safety,
- ❑ reliability

Key areas of activity

- Intellectual control and management systems
- Transportation management systems and transport service technology
- Signalling and remote control systems
- Automated transportation management centers
- Railway transport information systems
- Geoinformation systems and satellite technology
- Transport safety systems
- Infrastructure management systems
- Power consumption and energy management systems
- Testing, certification and expert assessment
- Information security
- Regulatory support

www.vniias.ru